

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:44:43 ; Search time 4729.08 Seconds
(without alignments)
10669.751 Million cell updates/sec

Title: US-09-913-767-10

Perfect score: 1067
Sequence: 1 agacaagaatgtgaaggt.....gaccataagtgaagttaa 1067

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1_0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_com.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1060.6	99.4	1067	6	AX033553 Sequence
2	977.6	91.6	121720	6	AC021044 Arabidops
3	476.4	44.6	1049	6	AX033545 Sequence
4	447.6	41.9	1075	6	AF078532 Arabidops
5	430	40.3	1250	8	AK117664 Arabidops
6	428.4	40.1	1102	8	BT005504 Arabidops
7	428.4	40.1	1225	6	AX033544 Sequence
8	428.4	40.1	1227	8	AF078531 Arabidops
9	330	30.9	114144	8	U78721 Arabidops
10	234.8	22.0	1458	8	AK061667 Oryza sat
11	230	21.6	1505	8	AK066610 Oryza sat
12	190.6	17.9	1363	8	AK064135 Oryza sat
13	182.2	17.1	1476	8	AK071254 Oryza sat
14	182.2	17.1	1617	8	AK072660 Oryza sat
15	182.2	17.1	175645	8	AP004096 Oryza sat
16	150	14.1	110036	8	AC142095 Medicago
17	147	13.8	1071	6	AX033550 Sequence
18	145.6	13.6	1165	8	AY096558 Arabidops
19	145.6	13.6	1428	8	AY074546 Arabidops

20	145.6	13.6	3387	6	AX507517 Sequence
21	145.6	13.6	3387	6	AX652015 Sequence
22	145.6	13.6	82697	8	ATT9A21 Arabidops
23	145.6	13.6	194143	8	ATCHRIV48 Arabidops
24	141.6	13.3	1152	6	AX653666 Sequence
25	141	13.2	14252	8	AF005399 Oryza sat
26	136.2	12.8	1194	6	AX033548 Sequence
27	135	12.7	1047	6	AX412868 Sequence
28	135	12.7	1081	6	AX033549 Sequence
29	135	12.7	1106	8	AF370622 Arabidops
30	133.6	12.5	73179	8	AP006383 Lotus cor
31	133	12.5	14894	8	AC121489 Oryza sat
32	133	12.5	195186	2	AC149309 Zea mays
33	131.4	12.3	1200	6	AX654080 Sequence
34	128.6	12.1	1506	8	AK105352 Oryza sat
35	127	11.9	96046	8	AC005967 Arabidops
36	119.8	11.2	1293	6	AX033547 Sequence
37	119.8	11.2	58427	8	ATF15J5 Arabidops
38	116.6	10.9	1145	6	AX033546 Sequence
39	114	10.7	1117	8	AY096468 Arabidops
40	114	10.7	1333	8	AY074535 Arabidops
41	114	10.7	83698	6	AX540653 Sequence
42	114	10.7	83698	8	AB010072 Arabidops
43	114	10.7	104386	8	ATT32A17 Arabidops
44	114	10.7	199362	8	ATCHRIV24 Arabidops
45	112.8	10.6	1251	8	BT003358 Arabidops

ALIGNMENTS

RESULT 1
AX033553
LOCUS AX033553 1067 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 10 from Patent WO0049152.
ACCESSION AX033553
VERSION AX033553.1 GI:10280295
KEYWORDS
SOURCE Arabidopsis thaliana (Chale cress)
ORGANISM Arabidopsis thaliana

REFERENCE 1
AUTHORS Andre B., Buerkle, L., Frommer, W.B. and Gillissen, B.
TITLE Nucleic acids that code for a nucleobase transporter
JOURNAL Patent: WO 0049152-A 10 24-AUG-2000;
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;
GILLISSEN BERND (DE)

FEATURES
source 1..1067
location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 99.4%; Score 1060.6; DB 6; Length 1067;
Best Local Similarity 99.6%; Pred No. 5.2e-269;
Matches 1063; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	1	AGACAAGAAATGGTGAAGGCTCTTGTGATCATAAACTGCATAATTCTACCCATAGGAAACT	60
Db	1	AGACAAGAAATGGTGAAGGCTCTTGTGATCATAAACTGCATAATTCTACCCATAGGAAACT	60
QY	61	GTGAGGTCCTTTGATATGCGTCTTACTTCAACAATGGCGGTAAGGATTTGGTTCT	120
Db	61	GTGAGGTCCTTTGATATGCGTCTTACTTCAACAATGGCGGTAAGGATTTGGTTCT	120
QY	121	CTAGTTTCTTGAACACTGCAGGCTTTCTGTTATCTTCTTCTCTCTCTCTCTTACA	180
Db	121	CTAGTTTCTTGAACACTGCAGGCTTTCTGTTATCTTCTCTCTCTCTCTCTTACA	180
QY	181	TTACCCGGCGCAAGCAACAATGTGGGTGATAGTACAAGTTTCTTCTTATCAACCCG	240

TITLE	Kawai, J., Hayaeshizaki, Y. and Shinozaki, K.
JOURNAL	Arabidopsis thaliana full-length cDNA
REFERENCE	Published Only in Database (2002)
AUTHORS	2 (bases 1 to 1250)
	Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayaeshizaki, Y. and Shinozaki, K.
TITLE	Direct Submission
JOURNAL	Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: maseki@gsc.riken.go.jp, URL: http://pfweb.gsc.riken.go.jp , Tel: 81-45-503-9625, Fax: 81-45-503-9586)
COMMENT	An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al. (1998) Plant J. 15:707-720; Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FIC-1-E vector (Carninci et al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://pfweb.gsc.riken.go.jp/) for further details.

```

FEATURES             Location/Qualifiers
     1..1250
         /organism="Arabidopsis thaliana"
         /mol_type="mRNA"
         /db_xref="taxon:3702"
         /chromosome="1"
         /clone="RAFL17-32-O17"
         /ecotype="Columbia"
         /note="common name: thale cress"
     1..1250
         /gene="Atlg28230/F3H9_10"
         53..1123
             /gene="Atlg28230/F3H9_10"
             /codon_start=1
             /product="putative purine permease"
             /protein_id="BAC42317.1"
             /db_xref="GI:26450405"
             /translations="MKNGLIINICIIITLTGCGPLLTRLTYFGSKRIWFMSEFLSTA
GPFIILIPLVSLFRRNRNPNNAENKTKLFLMETPLFTASIVIGLITGLDNYL
YSXGLAYLPVSTSSLIIGTQLAFNALFAFLVKQKTPPSINAVILTVGIGIALIHS
DGNAPFKSCKEIVGVGMVTAAALLYFAPILVELITYKKAROEITFPPLVLEIQMYMC
DKATFCVTGMFTVGDFKVIAREAREBFKGGSVFYALIVITGIWQGFPLGAIGIVF
CASLLASGVLIISLPVTEVPFAVAFFKPGVQAEKGVLSLLMGFVSFYFGPEPKSGIK
VVDKPQPPTLPILPVSDDYA"

```

ORIGIN		Query Match Best Local Similarity Matches 673; Conservative 0; Mismatches 315; Indels 24; Gaps 3;	40.3%; Score 430; DB 8; Length 1250; Pred. No. 1.7e-102;
QY	1 AGACAAGATGGTGAAGGCTCTTTGTGATCATATAAATGCATAAATCTTAGCCATAGGAAACT	60	
Db	45 AGAAGAAGATGAAGAATGGTTTGATAATCATAAATGTAATTATCCTCACATATAGGAACAT	104	
QY	61 GTGGAGGTCCTTTGATTAATATGCGTCTCTACTTCAACAATGGCGGTAAAGAAATTGGTTCT	120	
Db	105 GTGGAGGTCCTTTGTTAACTCGTCTCTACTTCCAAATGGCGGAAAACGAATCTGGTTCA	164	
QY	121 CTACGTTTTCTTGAAACTGACGGCTTTCCTGTTATCTTCATTCTCTGCTCTTCTCTTACA	180	
Db	165 TGAGCTTCTCTATCAACCGCTGGTTTTCCAATCATCTCATCCCTCTCTTGGTCTCCTCTCC	224	
QY	181 TTATCCGGCGCAACGAACAATGTGGTGA-----TAGTACAAAGTT	222	
Db	225 TCAGCCGTGCCCGAGAACCGCAACCTTACAACGGGAAAAACAAGCGAAACAAAGC	284	
QY	223 TCTTTCTTATCAAACCGCGCTCTTCTTATCGCCGCTGTATTGTAGGCATTCTCTCAGGGT	282	
Db	285 TCTTCTCATGGAACTCCTCTGTTTATCGCCTCCATTGTCATAGGGTTCCTCACAGGAC	344	
QY	283 TTGATAACTACTTGTATGCAATATGGTATAGCTTATCTTCCAGTTTCTACAGCTGCTCTTCA	342	

345	TGACAAC	TACTATATCTTACGGATTAGCAGATCTGCCAGTTTCAACTTTCATCGCTCA	404
343	TCATTGCTT	CACAGTTAGCTTTTATAGCTATCTTCTCATTTCTTCATGGTTAAACATAAAGT	402
405	TAATCGGAA	CTCAACTAGCTTCAACGCTCTCTTCGCTTCTCTGTTAGTCAAGCAAAAGT	464
403	TCACTCCTTTT	ACATCAATGCTGTTGTGTTTGTGACTGTTGGTGCCTGCGGTTTGGGAA	462
465	TCACTCGGTT	CTCATAAACCGCGTCGTTGTTGTGACGGTTGGTATCGGGATCCTTGCCT	524
463	TGCATACCGAAA	CTGATAAGCCAGTTCATGAGACTCAACGACGTACATAACTCGTTTCT	522
525	TACACATGAT	TGGACACAACCGGCTAAGAGACCAAGAAAGATATGTGGTTGGGTTCT	584
523	TGATTACTG	TAGCAGCAGCTGTTATGATGCTTTTCATCTTGCCATTAGTGGAACTTGCTT	582
585	TGATGACTG	TGGTTGTCAGCTCTTCTCTATGCTTTTATATTACCGCTCGTTGAGCTAACTT	644
583	ACCAGAAGCT	TAGCAACCATGAGCTATACCTTGTGCTCGAGTTCAGATGTGATTTGT	642
645	ACAAGAAGCT	CGTCAAGAAATCACTTCCACACTGTGTTCAGATTACAGATGTGCACT	704
643	GTCTCCTT	TGCTTATTTGTCAGCGCTCATCGGTATGTTTCATCGCTGGTGATTTCAAGCAGG	702
705	GCCTTGTCT	ACTTTTTCTGTGTCATTGGCATGTTTCATCGTTGGAGATTTTA--AGG	761
703	CCTTACCAA	AGAAGCAAGAGTTCAAGCTTGGAGAGCATGTGTT---CTATGTGGTGG	759
762	TGATAGCA	GAGAAGCAAGAGATTCAAGATTGGAGGATCAGTGTGTTTACTATGCATTGA	821
760	CTGTGTTTT	CAGCCATCATATGGCAAGGCTTCTTCTTGGGAGCCATTGGATTAATCTTCT	819
822	TAGTGATCA	CAGGAAATATATGGCAAGTTTCTTCTTAGGACCATAGGGAATCTGTTTT	881
820	CCACATCGT	CTCTCGTCTCGGTTATTATGATATCAGTGCCTTTTGCCAAATTACAGAGTTT	879
882	GTGCATCAT	CACTAGCTTCTGGTGTCTGTATAAGATGTCTCTGTTCCCGTGACTGAAGTTT	941
880	TAGCTGTAT	TATTTCTACCATGAAAGTTTCAAGCTGAAGGGAATTTCTTCTTGCTCTCT	939
942	TCGCGCTG	CTGTTGTTCCGGAGAGTTTCAGGCAGAGAAAGTGTCTCTACTTCTTT	1001
940	CCCTTTGG	CGCTTGTCTCTTACTTTTTATGGTGAGATAAAAGTCTGGCGAGGA	991
1002	CTCTTTGG	GAATGTTGTTCTTACTTCTACGGCAGTTTAAATCCGCAAGAA	1053

RESULT 6	BT0005504	1102 bp	mrna	linear	PLN 15-MAR-2003
LOCUS	BT0005504				
DEFINITION	Arabidopsis thaliana clone U50269 putative purine permease (At1g28230) mRNA, complete cds.				
VERSION	BT0005504	1	GI:28973198		
KEYWORDS	FLI CDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 1102)				
AUTHORS	Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.				
TITLE	Arabidopsis Open Reading Frame (ORF) Clones				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1102)				
AUTHORS	Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,				

Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.-J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (15-MAR-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, RGECC (SPD) Consortium members constructed and sequenced the PUNI (ORF) clones using the RFLV cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

FEATURES
source

Location/Qualifiers
1. .1102
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U50469"
/ecotype="Columbia"
/note="This clone is in pUNI 51."
1. .1102
/genes="Atlg28230"
1. .1071
/genes="Atlg28230"
/codon_start=1
/evidence=experimental
/product="putative purine permea
protein id: "AA063924.1"
/db_xref="GI:28971199"
/translation="MKNGLLINCIILITGFC
GFPIILPLVLFSLRNGSLRNPNNAENKRT
YSGYLAVPVSYSIIIGTQLAFNALFAFLV
DGNPKAKESKSSVVGFLMTVAALLVAFILP
LAATFCVIGMTIVGDGPKVIAAREEFKGG
CASSLACGSLVILPVPTEFAVVCREFKPOA
VMDKPPPTETLILPVPVSYYA"
1072. .1102
genes="Atlg28230"

gene

CDS

```

/genes="A1928230"
1. .1071
/genes="A1928230"
/codon_start=1
/evidence=experimental
/product="putative purine permease"
/protein_id="AAO63924.1"
/db_xref="GI:28973199"
/translation="MKGLGIIINCIITLTGTGCG
GFPIILPILSVFSSRRSRNPNNAENKRTK
YSLGVLAVPTSSLLITQLANFAFLVLYK
DGPDKAKESYVYGLMTVAALVYAFILPL
LAATFCVIGMFI VGDPKVTARAEFKIGSV
CASSLAGSVLIVLLPVPFAVVCREFQAE
VDKDPQSTPELIVLPVSDYVA"
1072. .1102
/genes="A1928230"

```

3' UTR

ORIGIN

Query Match 40.1%; Score 428.4; DB 8; Length 1102;
Best Local Similarity 66.6%; Pred. No. 4.6e-102;
Matches 669; Conservative 0; Mismatches 311; Indels 24;

Qy	9	ATGTGAAGGCTCTTGTGATCATAAACTGCATAATCTTAGCCATAGAAACCTGTGGAGGT	68
Db	1	ATGAAGAATGGTTTGATAATCATAAACTGTATTATCTCACTATAGAAACATGTGGAGGT	60
Qy	69	CCTTTGATTATGCGTCTCTACTTCACAAATGGCGGTAAAGGATTTGGTTCTCTACGTTT	128
Db	61	CCTTTGTTAAATCGTCTCTACTTCACCAATGGCGGAAAAACGAATCTGTTTCATGAGCTTC	120
Qy	129	CTTGAAACTCGAGCGTTTCTCGTTTATCTTCATTCCTCTGCTCTCTCTTACATTTACCGG	188
Db	121	CTATCAACCGGTGGTTTTTCCAATCATCTCATCCCTCTCTTGGTCTCTCTTCTCAGCGGT	180

Qy	189	CGCAGAGCAACAATGTGGGTGA-----TAGTACAAGTTCTTTCTTT	230
Db	181	CGCGCAGCAGCACCGCAACCCCTTAAACAACGCGGAAAAACAAGAGCTCTTCCTC	240
Qy	231	ATCAAAACGGGCTCTTCTTATCGCGCGTGTATGTAGGCAATCTCTCAGGGTTTGATAAC	290
Db	241	ATGGAACCTCTCTGTATTATCGGCTCCATTTGTATAGGGTGTCTCAGGACTTGACAAC	300
Qy	291	TACTTGTATGCATATGGTATAGCTTATCTTCCAGTTTTCTACAGCTGCTTTATCATTTGCT	350
Db	301	TACTTATATCTTACGGATTAGCATATCTGCCAGTTTCAACTTCATCGCTCATATCGGA	360
Qy	351	TCAGATTAGTTTTATAGCTATCTTCTCATTTCTCATGTGTTTAACATAAAGTTTCACTCT	410
Db	361	ACTCAACTAGCTTTTCAACGGCTCTCTTGGCTTTCTTGTAGTCAAGCAAAAAGTTTCACTCCG	420
Qy	411	TTTACCATCAATGCTGTGTGTGTGTGACATGTTGGTCTCGCGGTTTTGGGAATGCATACC	470
Db	421	TTTCCATAAACGCGCGTGTGTGTGACGGTGTGTATCGGGATCTTGGTGTACACAGT	480
Qy	471	GAAACTGATTAAGCCAGTTTATGAGACTTCAACAGCAGTACATAAATCGTTTCTTGAATTACT	530
Db	481	GATGGACAAACCGGCTAAGGAGAGCAAGAAGATATGTGGTGGTCTTGTATGACT	540
Qy	531	GTAGCAGAGCTGTATGTATGTCTTTCATCTTGCCATTAAGTGAACCTGTACCAGAAA	590
Db	541	GTGGTTCAGCTCTTCTCTATGTCTTTTATATATACCGCTCGTTGAGCTAACTTACAAGAAA	600
Qy	591	GCTAAGCAACCATGAGCTATACCCCTGTGCTCGAGTTCAGTTCAGTGTGTTGTGTCCTCTT	650
Db	601	GCTCGTCAAGAAATCACTTTCCCACTTGTGCTTGAGATTCAAGATGATGTGCTTGGCT	660
Qy	651	GCTTCTATTGTACGCGTCACTCGGTATGTTTCATCGCTGGTGTATTCAAGCAGGCTTACCA	710
Db	661	GCTACTTTTCTGTGTCATTTGGCATGTTTCATCGTTGGAGATTTTA---AGGTGATAGCA	717
Qy	711	AAAGAACCAAGAGATTCAAGCTTGGAGAGGCATTTGTT---CTATGTGGTGGCTGTGTTT	767
Db	718	AGAGAACCAAGAGATTCAAGATTGGAGGATCACTGTGTTTACTATGATTTGATGATGATC	777
Qy	768	TCAGCCATCATATGGCAAGGCTTCTTCTGGGAGCCATTTGGATTAATCTTCTCCACATCG	827
Db	778	ACAGGAATAATATGGCAAGGTTCTTCTTAGGAGCCATAGGGAATGTGTTTGTGCAATCA	837
Qy	828	TCCTCTGCTCGGATATTATGATATCAGTGTCTTTTGGCAATTAACAGAGGTTTTAGCTGTT	887
Db	838	TCAGTACTTCTGGTGTCTGATTAAGTGTCTCTCGGTGACTGAATTTTTCGCGCTC	897
Qy	888	ATATTTACCATGAAAGTTTCAAGCTGAGAGGGACTTTCTCTTGTCTCTCTCCCTTTGG	947
Db	898	GTTTGTGTTCCGGGAGAGTTTCAGGCAGAGAAAGTGTCTCTCTACTTCTTCTCTTTGG	957
Qy	948	GGCTTTGCTCTTACTTTTATGTTGATTAAGTCTTGGCGAGCA	991
Db	958	GGATTTGTCTTCTTACTTCTACGCGAGTTTAAATCCGCAAGAA	1001
RESULT 7			
AX033544			
LOCUS	AX033544	1225 bp	DNA
DEFINITION	Sequence 1 from Patent WO0049152.	linear	PAT 21-SEP-2000
ACCESSION	AX033544		
VERSION	AX033544.1	GI:10280288	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
REFERENCE	1		
AUTHORS	Andre, B., Buerkle, L., Frommer, W. B. and Gillissen, B.		
TITLE	Nucleic acids that code for a nucleobase transporter		
JOURNAL	Patent: WO 0049152-A 1 24-AUG-2000;		

RESULT 7
AY033544

AX033544	LOCUS	AX033544	1225 bp	DNA	linear	PAT 21-SEP-2000
	DEFINITION	Sequence 1 from Patent WO0049152.				
	ACCESSION	AX033544				
	VERSION	AX033544.1	GI:10280288			

SOURCE	ORGANISM	REFERENCE
Arabidopsis thaliana (thale cress)		
Arabidopsis thaliana		Andre, B., Buerkle, L., Frommer, W.B. and Gillissen, B.
Eubakryota		Nucleic acids that code for a nucleobase transporter
Karyokonta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		Patent: WO 0049153-A 1 24-AUG-2000.

FEATURES									
source									
LOCATION/Qualifiers									
1. .1225									
/organism="Arabidopsis thaliana"									
/mol_type="unassigned DNA"									
/db_xref="taxon:3702"									
ORIGIN									
Query Match 40.1%; Score 428.4; DB 6; Length 1225;									
Best Local Similarity 66.4%; Pred. No. 4.6e-102;									
Matches 672; Conservative 0; Mismatches 316; Indels 24; Gaps 3;									
Qy	1	AGACAAGATGGTGAAGCTCTTGTGATCATATAAAGTCAATAATCTAGCCATAGAAACT	60						
Db	17	AGAAGAAGATGAAGAATGGTTGATAATCATATAAAGTGTATTATCTCTACTATAGAACAT	76						
Qy	61	GTGAGAGTCCCTTGATTATGCGTCTCTACTTCAACAATGGCGGTAAGAAGATTGGTTCT	120						
Db	77	GTGAGAGTCCCTTGTTAACTGCTCTACTTCCACCAATGGCGGAACGAATCTGGTTCA	136						
Qy	121	CTACGTTTCTTGAAGTCAAGCTTTCCTGTATCTTCAATTCCTCTCTCTCTCTTACA	180						
Db	137	TGAGCTTCTATCAACCGCTGGTTTCCAAATCATCTCATCCCTCTCTTGGTCTCCTTCC	196						
Qy	181	TTACCGGGGAGAACCAATATGGGTGA-----TAGTACAAGTT	222						
Db	197	TGACCGGTGGCGGCAACCGAACCTTAACAACGCGGAACAAAGCGGAACAAAGC	256						
Qy	223	TCTTCTTATCAACCGGTCTCTTATCGCGGCTGTTATGTAGGCATTCTCTCAGGTT	282						
Db	257	TCTTCTCATGGAAGTCTCTGTTTATCGCTCTCATTTGTATAGGGTTGCTCAGGAC	316						
Qy	283	TGATAACTACTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	342						
Db	317	TTGACAACTACTTATTTCTTACGGATTAGCATATCTGCCAGTTTCAACTTCACTGCTCA	376						
Qy	343	TCATTGCTTCACAGTTAGCTTTATAGCTATCTTCTCTCTCTCTCTCTCTCTCTCTCT	402						
Db	377	TAATCGGAACCTCACTAGCTTTCAACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	436						
Qy	403	TCATCTCTTTTACCATCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	462						
Db	437	TCATCTCGGTTCTCCATTAACCGCGTCGTTTGTGTGACGGTTGGTATCGGATCCTTGGT	496						
Qy	463	TGATACCGAACTGATAGGCAAGTTATGATGATGATGATGATGATGATGATGATGATGAT	522						
Db	497	TACACAGTGTAGGACAAACCGGCTAAGGAGAGCAAGAAAGATGATGTGGTTGGTTCT	556						
Qy	523	TGATTTACTGTAGCAGCAGCTGTATGATGATGATGATGATGATGATGATGATGATGATG	582						
Db	557	TGATGACTGTGGTTGCGAGCTCTCTCTATGCTTTTATATACCGCTCGTGGAGTAACTT	616						
Qy	583	ACCAAGAGCTAAGCAAAACCACTAGCTATACCTTGTCTCGAGTTCAGATTGATTTTGT	642						
Db	617	ACAAGAAGCTGTCAGAAATCACTTTCCCACTTGTCTTGAGATTGAGATGTCATGT	676						
Qy	643	GTCTCCCTTCTTATGTCAGCGTCATCGGTATGTTTATGTTTATGTTTATGTTTATGTTT	702						
Db	677	GCTTGTCTGCTACTTTTCTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC	733						
Qy	703	CCTTACCAAAAGACGACGAGGTTCAAGCTTGGAGGCAATGTT---CTATGTGGTG	759						
Db	734	TGATAGCAAGAGAACGAGAGGTTCAAGATTGGAGGATCAGTGTTTTACTATGCAATGA	793						
Qy	760	CTGTGTTTTCAGCCATCATATGGCAAGGCTTCTTCTGGGAGCCATTGGATTAATCTTCT	819						
Db	794	TAGTATCACAGGAATATATGGCAAGGTTTCTTCTTAGGAGCCATAGGATTTGTTTT	853						
Qy	820	CCAATCGTCTCTCGGTATATGATATCATAGTCTTTTTCGCAATTCAGAGGTTT	879						
Db	854	GTGCATCATCACTAGCTTCTGGTGTCTGATAAGTGTCTGCTTCGGTACTGAAAGTTT	913						

Qy	880	TAGCTGTATATCTTACCATGAAAAGTTTCAAGCTGAGAGGAGCTTCTCTCTCTCTCT	939						
Db	914	TGCGCGTCGTTGTTTCCGGAGAGTTTCCGGAGAGAGGTTCTCTCTCTCTCTCTCT	973						
Qy	940	CCCTTTGGGGCTTGTCTCTTACTTTTATGGTGAGATAAAAGTCTGGCGAGGA	991						
Db	974	CTCTTTGGGATTTGTCTCTTACTTCTACGCGAGTTTAAATCCGGCAAGAA	1025						
RESULT 8									
AF078531									
LOCUS									
Arabidopsis thaliana 1227 bp mRNA linear PLN 20-APR-2000									
ACCESSION									
AF078531									
VERSION									
AF078531.1 GI:7620006									
KEYWORDS									
SOURCE									
Arabidopsis thaliana (thale cress)									
ORGANISM									
Arabidopsis thaliana									
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;									
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;									
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi									
REFERENCE									
AUTHORS									
1 (bases 1 to 1227)									
Gillissen,B., Burkle,L., Andre,B., Kuhn,C., Rentsch,D., Brandl,B.									
and Frommer,W.B.									
TITLE									
A new family of high-affinity transporters for adenine, cytosine,									
and purine derivatives in Arabidopsis									
JOURNAL									
MEDLINE									
PUBMED									
20129770									
REFERENCE									
AUTHORS									
2 (bases 1 to 1227)									
Gillissen,B., Andre,B., Rentsch,D., Buerkle,L., Kuehn,C., Brandl,B.									
and Frommer,W.B.									
TITLE									
Direct Submission									
JOURNAL									
Submitted (15-JUL-1998) Institut fuer Botanik, University of									
Tuebingen, Auf der Morgenstelle 1, Tuebingen, Baden-Wuerttemberg									
72076, Germany									
FEATURES									
Location/Qualifiers									
1. .1227									
/organism="Arabidopsis thaliana"									
/mol_type="mRNA"									
/cultivar="Columbia"									
/db_xref="taxon:3702"									
1. .1227									
/gene="PUP1"									
27. .1097									
/gene="PUP1"									
/note="AtPup1"									
/codon_start=1									
/product="purine permease"									
/protein_id="AA564547.1"									
/db_xref="GI:7620007"									
/translation="MNGLIIINCIILITGTCGGLLRLTYFNGKRIWFMFLSTA									
GFPIILPLVSLRRGNPNNAENRKLFLMETPLFIASIVIGLTLGLDNL									
YSYGLAYLVSTSLIIGTQAFNALFAFLVKQKTFPSINAVLTVGIGIALHS									
LQATPKESKKEVVGFLMTVAALYAFLLPLVELTYKKAROEITPLVLEIQMVC									
DGATFCVIGMFIVGPKVIAREAREFKIGSVFYVALIVITGIMQGFILGIVF									
CASLASGVLSVLLPVTVEFVAVCVREKQAEKVSLLLSLGMFVSFYGFEPKSGKK									
VVDKPPFPPELPLPVSDIYA"									
ORIGIN									
Query Match 40.1%; Score 428.4; DB 8; Length 1227;									
Best Local Similarity 66.4%; Pred. No. 4.6e-102;									
Matches 672; Conservative 0; Mismatches 316; Indels 24; Gaps 3;									
Qy	1	AGACAAGATGGTGAAGCTCTTGTGATCATATAAAGTCAATAATCTAGCCATAGAAACT	60						
Db	19	AGAAGAAGATGAAGAATGGTTGATAATCATATAAAGTGTATTATCTCTACTATAGAACAT	78						
Qy	61	GTGAGAGTCCCTTGATTATGCGTCTCTACTTCAACAATGGCGGTAAGAAGATTGGTTCT	120						
Db	79	GTGAGAGTCCCTTGTTAACTGCTCTTACTTCCACCAATGGCGGAACGAATCTGGTTCA	138						
Qy	121	CTACGTTTCTTGAAGTCAAGCTTTCCTGTATCTTCAATTCCTCTCTCTCTCTTACA	180						

Db 25959 GCTTCAGTCTCTCTTTGCACTTTTATGCTGAAGCAAAAGTTTCACACCTTTTCACTATAA 25900
 QY 421 ATGCTGTGTGTTGTTGACTGTTGCTGCGGTTTGGGAATGCATACCGAAACGTGATA 480
 Db 25899 ACCTATCGTTTGTCTCAGTGTGGTCCGTAGTCTTGCCTTAATCTCTGATAGTGACA 25840
 QY 481 AGCCAGTTCATGAGACTCACAACAGTACATACTGGTTTCTTGATTACTGTAGCAGCAG 540
 Db 25839 AGCTTGCAACAGAGACACACAAGGAATATGTTGGGTTTCATCATGACTCTTGGTGCAG 25780
 QY 541 CTGTTATGATGCTTTTCATCTTCCATAGTGAACCTGCTTACCAAGAAAGCTAACCAAA 600
 Db 25779 CTCTTCTATGGTTTATATATGCCACTTGTGCAGCTTCTTTACAAGAAATCTGGTCAGC 25720
 QY 601 CCATGAGCTATACCTTGTGCTGAGTTCAGTTGATTTGTTGCTCTCTTCTTCATTG 660
 Db 25719 GAATCAGTATACGTCGCGCTCGAGTTCAGATGCTTATGCTTGTGCCCACTGTG 25660
 QY 661 TCAGCGTCATCGGTATGTTTCATCGCTGGTGATTTCAG 698
 Db 25659 TCTGCTCGTGGGATGCTAGCGCTGGCGATTTCAG 25622

RESULT 10

AK061667 1458 bp mRNA linear PLN 24-JUL-2003
 LOCUS Orzyza sativa (japonica cultivar-group) cDNA clone:001-036-F06, full
 DEFINITION insert sequence.
 ACCESSION AK061667
 VERSION AK061667.1 GI:32971695
 KEYWORDS FLI CDNA; oligo-capping.
 SOURCE Orzyza sativa (japonica cultivar-group)
 ORGANISM Orzyza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

1 The Rice Full-Length cDNA Consortium, National Institute of
 Agricultural Sciences Rice Full-Length cDNA Project Team:
 Kikuchi,S., Satoh,K., Negata,T., Kawagashira,N., Doi,K.,
 Kishimoto,N., Yazaki,K., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
 Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
 Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group: Ootomo,Y., Murakami,K.,
 Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
 Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
 Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
 Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
 Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,
 Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
 Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
 Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
 Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
 Yoshino,M. and Hayashizaki,Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from
 japonica rice

Science 301 (5631), 376-379 (2003)
 22752273
 12869764
 2 (bases 1 to 1458)
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
 Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
 Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
 Hori,F., Hotta,I., Iida,Y., Ikeda,R., Imamura,K., Imotani,K.,
 Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
 Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
 Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
 Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
 Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
 Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
 Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
 Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,

TITLE
JOURNAL

COMMENT

FEATURES

source

1. 1458
 organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiyar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="001-036-F06"

ORIGIN

Query Match 22.0%; Score 234.8; DB 8; Length 1458;
 Best Local Similarity 54.7%; Pred. No. 6.1e-51;
 Matches 542; Conservative 0; Mismatches 427; Indels 21; Gaps 3;
 QY 21 CTTGTGATCATAACTGCATAATTTAGCCATAGGAACTGTGGAGTCTCTTTGATTATG 80
 Db 204 CTCGTGGCGCTCACTGCGGATGCTGGTGGGACCACCGGGGGCGATCCCTAGC 263
 QY 81 CGTCTCTACTTCAACAATGGCGGTAAAAGGATTTGTTCTCTTACGTTTCTGAAACTGCA 140
 Db 264 CGGCTCTACTTCAGCAAGGGCGGCACAGGAAGTGGCTCTCGGCTGGGTCTGAGACCGCC 323
 QY 141 GCGTTTCTGTGTTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 200
 Db 324 GGGTGGCGGCT 383
 QY 201 AATGTGGGTGATAGTACAAAGTTTCTTCTTATCAAAACGCGTCTTCTTATCGCGCTGT 260
 Db 384 CCGC-----GGCGCGCGTGTCTCTACGCGCGCGGCTGCTGCTGGCGGCGG 434
 QY 261 ATTGTAGGCAATTCCTCAGGGTTTGATAACTACTTGTATGTCATATGTTATGTTATCTT 320

Direct Submission
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
 Agricultural Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
 305-8602, Japan [E-mail:skikuchi@nias.affrc.go.jp,
 Tel:81-29-838-7007, Fax:81-29-838-7007]
 This clone is one of the 28K full-length cDNA clones from japonica
 rice.

URL : http://cdna01.dna.affrc.go.jp/cDNA/
 NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
 Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
 Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
 Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
 Yamamoto,M.

FAIS Genome Sequencing & Analysis Group: Ootomo,Y., Iida,Y.,
 Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
 Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
 Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
 Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
 Yoshimura,A., Matsubara,K. and Murakami,K.
 Genome Exploration Research Group in Riken Genomic Sciences Center
 and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
 Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
 Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
 Hiroka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
 Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
 Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
 Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
 Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
 Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
 Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,A., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
 Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toyota,T., Waki,K.,
 Yasunishi,A. and Hayashizaki,Y.
 Location/Qualifiers
 1. 1458

```
Db      435  GTGCTCGGGTCTCTACCGCGCGCGACGACTTCGTCTACGCTCAGCGCTCGCCCTACCTG 494
Qy      321  CAGATTTCACAGCTGCTCTATCATCTGCTTCACAGTTAGCTTTTATAGCTATCTCTCA 380
Db      495  CGGTGTCCACCTCGCCCATCTCTCATCTCCACGAGCTCGCGTTCCACCGTCTTCTTCGG 554
Qy      381  TTCTTCATGGTTAAACATAAGTTCACTCTCTTTTACCATAAAGTCTGTTGTGTTGTGACT 440
Db      555  TGCCTCATGTCGGCGAGCGCTCACCGGCGCGACGCTGAACGCGTGGCGTGTCTCACT 614
Qy      441  GTTGTGTCGGGTTTGGGAATGCATACCGAAACTGATAAGCCAGTTCATAGAGACTCAC 500
Db      615  ATCGCGCGCGTGTGTGGCGCTCCACGCGTCCAAAGGACCGACCGCGCGGGTGACCAACC 674
Qy      501  AAGCAGTACATAACTGGTTCTTGTGATTACTGTAGCAGCAGCTGTATGTATGCTTTGATC 560
Db      675  GGAAGATCTGGATGGGCTTCTCTCTCACCTCGCGCGCGCGCTGTACGGGCTCATC 734
Qy      561  TTGCCATTAGTGGAACTTGCTTTACCAGAAAGCTAAG-----CAAACCATGAGCTAT 611
Db      735  CTGCGCTGCTGAGCTCGGTACAGCAGCGCGCGCGCGCGCGCGCTGCTACGTAC 794
Qy      612  ACCCTTGCTCGAGTTCCAGTTGATTTTGTGTTCTCTTGTCTTCTTATTTGTGAGCTCATC 671
Db      795  GCGCTGGTGATGGAGATGAGCTGCTCATGGCTTCTTTCGCCACCGCGCTTCTGTACCGCTC 854
Qy      672  GGTATGTTTCATCGCTGGTGTATTTCAAGCAGGCTTACCAGAGAGCTTACCAGAGAGAGTTCAG 731
Db      855  GGTGTGATGCTGTAACAGAGCTTC---CAGGCGATCCCGAGGAGCGCAAGCAGTACGAG 911
Qy      732  CTTGGAGAGGCAATGTTCTTATGTGCTGTGTGTTTTCAGGCATCATATGCAAGGCTTC 791
Db      912  CTGGGTGAAGCGCGTACTAGCTGTGTGTGTTTCAACGCGCTCTCTGGAGTTCCTTC 971
Qy      792  TTCCTGGAGGCAATGATTTCTTCTCCACATCGTCTCTGCTCGGTCCTGGGTATTATGATA 851
Db      972  TCTGTGGGCGCGTGGCGTCTATCTTCTGCTGCACAGCGCTGCTCGCGGATCATCATC 1031
Qy      852  TCAGTGTCTTTTGCCAAATACAGAGTTTATGCTGTATATCTACCATGAAGTTTCAA 911
Db      1032  GCGGTGTTATCCCGATCAGGAGTGCTGGCGTCACTTCCTCCACGAGAGTTACG 1091
Qy      912  GCTGAGAAAGGCACTTCTCTGCTCTCTCCCTTTGGGCTTGTCTCTTACTTTTATGGT 971
Db      1092  AGCGAGAAAGGCGTGGCGCTGTGCTCTGCTCTGGGCGCTCGCTCTCTCTCTACGCG 1151
Qy      972  GAGATAAAGTCTGGCGAGGATAAAGGAGA 1001
Db      1152  GAGTACGCGCGAGCCCAAGCCCAAGAGAAA 1181
```

```
RESULT 11
AK066610
LOCUS   AK066610      1505 bp      mRNA      linear      PLN 24-JUL-2003
DEFINITION  Oryza sativa (japonica cultivar-group) cDNA clone:J013073M15, full
            insert sequence.
ACCESSION AK066610
VERSION   AK066610.1 GI:32976628
KEYWORDS  FLI_CDNA; CAP trapper.
SOURCE    Oryza sativa (japonica cultivar-group)
           Oryza sativa (japonica cultivar-group)
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1
AUTHORS   The Rice Full-Length cDNA Consortium, National Institute of
            Agrobiological Sciences Rice Full-length cDNA Project Team,
            Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
            Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Hotta,I.,
            Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
            Ohtsuki,K., Shishiki,T., and
            Science Genome Sequencing & Analysis Group; Otomo,Y., Iida,Y.,
            Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
            Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
            Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
            Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
            Yamamoto,M.
```

```
TITLE     Direct Submission
JOURNAL    Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
MEDLINE    Agrobiological Sciences, Department of Molecular Genetics, Head of
PUBMED     Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
REFERENCE  305-8602, Japan (E-mail:ekikuchi@nias.affrc.go.jp).
           Tel:81-29-838-7007, Fax:81-29-838-7007
COMMENT    This clone is one of the 28K full-length cDNA clones from japonica
           rice.
           URL : http://cdna01.dna.affrc.go.jp/cDNA/
           NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,
           Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
           Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
           Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
           Yamamoto,M.
           PAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
           Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
           Nagata,T., Kuroosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
           Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
           Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
           Yoshimura,A., Matsubara,K. and Murakami,K.
           Genome Exploration Research Group in Riken Genomic Sciences Center
           and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
           Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
           Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
           Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
           Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
           Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
           Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
           Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
           Ota,Y., Saichoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
           Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
           Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
           Takaku-Akaiura,S., Tanaka,T., Tomaru,A., Toyota,T., Waki,K.,
           Yasunishi,A. and Hayashizaki,Y.
           Location/Qualifiers
               1. .1505
```

Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashizume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Sato,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

2 (bases 1 to 1505)

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori.F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kogawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saichoh,H., Sakai,C., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akaiura,S., Tanaka,T., Tomaru,A.,
Toyota,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
Yoshimura,A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:ekikuchi@nias.affrc.go.jp).

Tel:81-29-838-7007, Fax:81-29-838-7007

This clone is one of the 28K full-length cDNA clones from japonica
rice.

URL : http://cdna01.dna.affrc.go.jp/cDNA/

NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,
Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
Yamamoto,M.

PAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Nagata,T., Kuroosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
Ota,Y., Saichoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
Takaku-Akaiura,S., Tanaka,T., Tomaru,A., Toyota,T., Waki,K.,
Yasunishi,A. and Hayashizaki,Y.

Location/Qualifiers

1. .1505

FEATURES

source

Kojima, K., Naniki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Otsu, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

2752273

12869764

2 (bases 1 to 1617)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Horii, F., Hotta, I., Iida, J., Iishi, Y., Ishikawa, M., Itoh, M., Kagawa, I., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kameda, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Naniki, T., Narikawa, R., Nikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Otsu, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugawana, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Naniki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Horii, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kameda, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,

Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1..1617
/organism="Oryza sativa (japonica cultivar-group)"
/molecule="mRNA"
/db_xref="taxon:39947"
/clone="J023134011"

ORIGIN

Query Match 17.1%; Score 182.2; DB 8; Length 1617;
Best Local Similarity 50.9%; Pred. No. 4.7e-37;
Matches 458; Conservative 0; Mismatches 438; Indels 3; Gaps 1;

QY 63 GGAGGTCCTTGTGATATGCGTCTCTACTTCAACAATGGCGGTAAAGGATTGGTTCTCT 122
DB 459 GCATCTACACTTCTTGGGAGGTTCTACTACAATCAAGGCGGCAATAGCAAGTGGATGCC 518
QY 123 ACGTTTCTTGAACCTGCGAGGCTTTCTGTATCTTCTATCTCTCTCTCTCTCTCTTACATT 182
DB 519 ACATTGCTCCAACTGCTGGCTTTCCGATTTTGTTCATTGCCCCTATTTCTTTTCCATTCA 578
QY 183 ACCCGGCGCAGAACCAACAATGTGGGTGATAGTACAAGTTTCTTTCTTATCAAAACCGGT 242
DB 579 AAGACATCTTCTACACAAACAGTCACTAGTAGTCTCTGCCCTACAAATTTCTATCCCAA 638
QY 243 CTTCTTATGCGCGCTGTATGTAGGATCTCTCAGGTTTGTATACTACTTGTATGCA 302
DB 639 ATTACTCTGATATATGTGCTGGGCTCATCATTTCTGCAGACGACTTGTATTTCC 698
QY 303 TATGCTATAGCTTATCTTCCAGTTTCTACAGTCTCTCTTATCATTTCTTACAGTTAGCT 362
DB 599 TATGCGCTACTATATCTTCGGTCTCAACATATTCGCTCATCTGTGTAGTCACTTGGCC 758
QY 363 TTTATAGCTATCTTCTCTCATTTCTCATGGTTTAAACATAAAGTTTCACTCTTTTACCATCAAT 422
DB 759 TTCAATGCTGCTCTCTCATATTTCTCAATGCTCAAAAATTCACCCCTCTGATTTTCAAT 818
QY 423 GCTGTTGTTGTGACTGTGGTCTGGGTTTGGGAATGTCATCCGAACTGATAAG 482
DB 819 TCCGCTAGTCTCTTACGTTTCTGCTTCACTCTCTGGAGTTGATGAAGATTCTTCAGGGA 878
QY 483 CCAGTTTATGAGACTCAACAGCAGTACATAAAGTGTCTTCTGATTCTGTAGCAGCAGCT 542
DB 879 ACTACTAGTATATACATGGAGTACATTTTGGTTTCTGTGTGACACTAGGGGCATCA 938
QY 543 GTTATGTATGCTTTTCACTTGTGCCATAGTGAACCTGCTTACCAAGAAAGCTTAAGCAACC 602
DB 939 GCTACATCTCGCTCACTTCTCTCCCTGATGCAAGTCAATTTGAGAAGGTTATTAAGAGG 998
QY 603 ATGAGCTATACCTTGTGCTCGAGTTTCCAGTTGATTTTGTGCTCTCTCTCTCTCTCTTATGTC 662
DB 999 GAGACCTTCTCAGTTGTGTGAACATGATATATACAGCTCTCTGTTGGGCAACATTTGGCT 1058
QY 663 AGCGTCTATCGTATGTTTCACTCGCTGTGTTTCAAGCAGGCTTACCACAAAGCAAGA 722
DB 1059 TCTCTTGTGGTTTATTTGCAAGTGGTGAATGGATGA--CTTTACAGAGGAGATGCAT 1115
QY 723 GAGTTCAAGCTTGAGAGGAGCATTTTCTATGTGTGTGCTGTGTTTTCAGGCATCATATGG 782
DB 1116 GCATTCCAGTCTGGGAAGCTGTATATGTAATGACACTGCTGTGGAGGCTATATCTTGG 1175
QY 783 CAGGCTTCTTCTTGGGAGGCAATTTGATTAATCTTCTCCACATCGTCTCTCGTCTCGGT 842
DB 1176 CAGTATGATATCAGTTGGGAGTGGTGGGATGATCTTTGTGGTGCATCGCTGTTTTCAAAT 1235
QY 843 ATTATGATATCAGTCTCTTTTGGCCAAATTTACAGAGTTTGTAGCTGTTATATTTCTACCATGAA 902
DB 1236 GTGATAAGCACCTAGCTCTACCCATCATTTCTCTGTTTGTGTGATTTTCTTTTCATGAC 1295
QY 903 AAGTTTCAAGCTGAGAAAGGAGCTTTTCTTGTGCTCTCTCTCTCTCTCTTGTGGGCTTTGTCTCTTA 961
DB 1296 AAGATGATGAGGTAAGATTATAGCTATCTGATGGCCATTTGGGAGTTTATGTCATA 1354

CDS
/gene="OJ1743_B12.6"
/note="start and end point are not identified"
complement(join(30415..30625,30657..30815,30939..31122,
31157..31265))
/gene="OJ1743_B12.6"
/note="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD07657.1"
/db_xref="GI:41052789"
/translation="MGQTPRKWRORRPPPLTFQWTVHPIELWSHFGVSGGLARRVP
PKRNYGTVANSLSLSPKSGDVAENMTSLNLTLELTLYVTPVPPFPRGRGAAY
RLHRLGSDIVTGHSDVDAASELLPSTESVNTSLGASLTFRPAAVTERAAGEAGR
GGCRARWRASAAAGERGCGGCRFEVAQQLQRRPPLAPLRPRISACFARAPVLAPSL
TR"

gene
complement(31384..31803)
/gene="OJ1743_B12.7"
complement(<31384..>31803)
/gene="OJ1743_B12.7"

CDS
/note="start and end point are not identified"
complement(31384..31803)
/gene="OJ1743_B12.7"
/note="predicted by GlimmerM etc."
/codon_start=1

gene
/product="hypothetical protein"
/protein_id="BAD07658.1"
/db_xref="GI:41052790"
/translation="MAAAGGHSGDARPGGAPRRGRAATLGRROGGAGVCRRRGR
LEGADAGIDRARRRRTAARRPAVGGHEGQAPSPSSRRREVLTDASSWCCS
FSQDLPLVLHYTEQIATVLTQNRASRPWEEGK"
complement(32508..32834)
/gene="OJ1743_B12.8"
complement(<32508..>32834)
/gene="OJ1743_B12.8"

CDS
/note="start and end point are not identified"
complement(32508..32834)
/gene="OJ1743_B12.8"

/note="predicted by GeneMark.hmm etc."

gene
/product="hypothetical protein"
/protein_id="BAD07659.1"
/db_xref="GI:41052791"
/translation="MALIWNLPFARNTRMTTKEHQHRYGQDARGDDDKWHALTPFLV
AASCTVAPSVRHGGVPPSVVDSTSGVREPCVLVRWQVTSQRRRDTTSMAMPMLLPP
IQATTS"
3135..34677
/gene="OJ1743_B12.9"

join(<3135..33296,34573..34677)

gene
/note="start and end point are not identified"
join(3135..33296,34573..34677)
/gene="OJ1743_B12.9"

CDS
/note="predicted by GlimmerM etc."

gene
/product="hypothetical protein"
/protein_id="BAD07660.1"
/db_xref="GI:41052792"
/translation="WACHRHFPVQDVQASPLLQISLQKAGERNRKKEGEREGGR
17.1%; Score 182.2; DB 8; Length 175645;
Best Local Similarity 50.9%; Pred. No. 5.3e-37;
Matches 458; Conservative 0; Mismatches 438; Indels 3; Gaps 1;

Query Match
Best Local Similarity 50.9%; Pred. No. 5.3e-37;
Matches 458; Conservative 0; Mismatches 438; Indels 3; Gaps 1;

OJ 63 GGAGGTCTTTGATTATCGTCTACTTACACATGGCGGTAAAGGATTGGTTCTCT 122
Db GCATCTACACTCTTTGGGAGGTTCTACTACATCAAGCGGCAATAGCAAGTGGATGCC 78763

OJ 123 ACGTTTCTTGAACGACGAGGTTTCCTGTTATCTTCTTCTCTCTCTCTCTCTCTCT 182

Db ACATTGCTCCAAACGCTGGCTTTCGGATTGTTTCATTCGCTTCTCTCTCTCTCTCT 78764

OJ 183 ACCCGGCGCAGACGACCAATGTGGGTGATAGTACAACTTTCTTCTTCTCTCTCT 242

Db 78824 AAGACATCTTCTACACAAAACAGTCACTAGTAGTCTCTGCCCCCTACAAATTTCTATCCCCAAA 78883
OJ 243 CTTCTTATCGCGGCTGTTTATTTAGTGGCATTTCTCTCAGGGTTTGTAACTACTACTTGTATGCA 302
Db 78884 ATTACTCTGATATATGTTGCTCTGGGCTCATCAATGCTGCAGACGACTTGATGTATTTCC 78943
OJ 303 TATGGTATAGCTTATCTTCCAGTTTCTACAGCTGCTCTTATCATTTGCTTTCACAGTTAGCT 362
Db 78944 TATGGCTACTATATCTTCGGTCTCAACATATTTGGCTCATCTGTGTAGTACAGTTGCTGCC 79003
OJ 363 TTTATAGCTATCTTCTCATCTTTCATGTTTAAATAAAGTTTCACTCTCTTTTACCATCAAT 422
Db 79004 TTCAATGCTGCTTCTCATATTTCTCAATGCTCAAAAATTCACCCCTCTGATTTTCAAT 79063
OJ 423 GCTGTTGTTGTTGACTGTTGTTGGGCTGCTCGGTTTGGGAATGCATACCGAACTGATAG 482
Db 79064 TCCGTAGTCTCTCTTACGTTTCTGCTTCACTCTCTGGAGTTGATGAAGATTTCTCAGGGA 79123
OJ 483 CCAGTTTATGAGACTCAACAGCAGTACATAACTGTTTCTTGTATTCTGTAGCAGCAGCT 542
Db 79124 ACTACTAGTATATCACATGGGAAGTACATTTTGGGTTTCTGTGTACACTAGGGGCATCA 79183
OJ 543 GTTATGATGCTTTTCACTCTGCCAATAGTGAACCTGCTTACCAAGAAAGCTTAAGCAAAACC 602
Db 79184 GCTACATACTCGCTCAITTTCTCCCTGATGCAAGTCAATTTGAGAAGGTTATTAAAGAGG 79243
OJ 603 ATGAGCTATACCTTGTGCTCGAGTTCCAGTTGATTTTGTGCTCTCTCTCTCTCTCTCTCT 662
Db 79244 GAGACCTTCTCAGTTGTTTGAACATGCAGATATATACAGCTCTCTGTTGGCAACATTTGGCT 79303
OJ 663 AGCGTCACTCGGTATGTTTCACTCGCTGGTGAATTTCAAGCAGGCTTTACCAAAAAGAACAGA 722
Db 79304 TCTCTTGTGGGTTATTTGCAAGTGGTGAATGGATGA---CTTTACAAGGAGAGATGCAT 79360
OJ 723 GAGTTCAAGCTTGAGAGGCAATGTTCTATGTGTGGTGGTGTGTTTTCAGGCAATCATATGG 782
Db 79361 GCATTCAGTCTGGAAGCTGTATATGTAATGACACTGCTGTGGACGGCTATATCTTGG 79420
OJ 783 CAAGGCTTCTTCTGGGAGCCATTTGAATTAATCTTCTCCACATGCTCTCTGCTCTCGGT 842
Db 79421 CAGGTAGCATCAGTTGGAGTGGTGGGATGATCTTTGTGTTGTCATCGCTGTTTTCAAAT 79480
OJ 843 ATTATGATATCAGTGTCTTTTGGCCAAATTAACAGAGGTTTATAGCTGTTATATTTCTACCAATGAA 902
Db 79481 GTGATTAAGCACCTTAGCTCTACCCATCATTTCTCTGTTTGTGTGATTTTCTTTTCATGAC 79540
OJ 903 AAGTTTCAAGCTGAGAGGGAATTTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 961
Db 79541 AAGATGGATGGAGTAAAGATTTATAGCTATGCTGATGGCCATTTGGGGATTTTATGTCATA 79599

Search completed: November 1, 2004, 15:44:53
Job time : 4734.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:40:28 ; Search time 626.985 Seconds
(without alignments)
9996.736 Million cell updates/sec

Title: US-09-913-767-5
Perfect score: 1194
Sequence: 1 tcatgagatataataaacat.....tgtgatcaagcatatttcc 1194

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues 8269772
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1190.8	99.7	1194	3	Aaa97923 A. thalia
2	1015.8	85.1	3387	6	Abz14407 Arabidops
3	1015.8	85.1	3387	8	Ada68246 Arabidops
4	813.4	68.1	1293	3	Aaa97922 A. thalia
5	459.8	38.5	592	10	Abx56692 Arabidops
6	405.8	34.0	1071	3	Aaa97925 A. thalia
7	369.6	31.0	1047	6	Adg88190 A. thalia
8	369.6	31.0	1081	3	Aaa97924 A. thalia
9	317.2	26.6	1145	3	Aaa97921 A. thalia
10	316.8	26.5	1292	3	Aac51119 Arabidops
11	316.8	26.5	1295	3	Aac37289 Arabidops
12	305.8	25.6	492	3	Aac36845 Arabidops
13	270	22.6	2660	3	Aac47974 Arabidops
14	215.4	18.0	83698	6	Abn85767 Arabidops
15	213.2	17.9	2175	8	Ada70748 Rice gene
16	212.8	17.8	1152	8	Ada70213 Rice gene
17	196	16.4	1403	3	Aac37652 Arabidops
18	135.4	11.3	1049	3	Aaa97920 A. thalia
19	134.4	11.3	1155	8	Ada69686 Rice gene
20	120	10.1	386	6	Abq85215 Arabidops
21	118.4	9.9	332	12	Adp92185 Cotton ex

22	111.8	9.4	1225	3	Aaa97919
23	111	9.3	2586	3	Aac44184 Arabidops
24	103.2	8.6	1040	8	Ada69568 Rice gene
25	75	6.3	5520	5	Aac68924 FLR1 rece
26	73.2	6.1	1173	12	ADN72520
27	66.2	5.5	446	3	AAC37188
28	65.8	5.5	1421	3	AAC39779
29	64.6	5.4	477	3	AAC37273
30	64.2	5.4	1418	3	AAC46395
31	62.6	5.2	774	6	ABN99127
32	51.6	4.3	1200	8	ADA70627
33	50.6	4.2	2000	8	ADA71938
34	43.6	3.7	493	3	AAC36831
35	42.2	3.5	2000	8	ADA71938
36	40.4	3.4	349901	10	ADC86940
37	37.4	3.1	358	6	ABV98933
38	36.8	3.1	369	4	Aak80520 Human imm
39	36.8	3.1	369	4	Aak80518 Human imm
40	36.8	3.1	369	4	Aak80519 Human imm
41	36.8	3.1	370	4	Aak61600 Human imm
42	36.8	3.1	6741	3	AAA10595 Gene enco
43	36.4	3.0	651	10	ADF57711 Human pol
44	36.4	3.0	2928	3	AAC39967 Arabidops
45	36.4	3.0	4590	5	Aah24065 Yeast AOD

ALIGNMENTS

RESULT 1
AAA97923
ID AAA97923 standard; DNA; 1194 BP.
AC AAA97923;
XX

DT 19-JAN-2001 (first entry)
XX

DE A. thaliana PUP1 DNA #5.
XX

KW PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.
XX

OS Arabidopsis thaliana.
XX

PN DE19907209-A1.
XX

PD 24-AUG-2000.
XX

PF 19-FEB-1999; 99DE-01007209.
XX

PR 19-FEB-1999; 99DE-01007209.
XX

PA (FROM/) FROMMER W.
XX

PI Gillissen B, Buerkle L, Andre B, Frommer WB;
XX

DR WPI; 2000-566202/53.
XX

PT Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transporter, encodes a nucleobase transporter protein of
Arabidopsis thaliana.
XX

PS Claim 1f; Page 14; 24pp; German.
XX

CC This invention describes a novel nucleic acid encoding a plant nucleobase
CC transporter (I). (I) is produced by complementation of a nucleobase
CC transport (NBT)-defective host cell with a plant gene bank by selection
CC of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
CC expression of (II) (when in antisense orientation), and to produce
CC transgenic crop plants. The transgenic plants have modified nucleobase
CC transport properties, e.g. altered affinity and substrate specificity

CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX
SQ Sequence 1194 BP; 298 A; 279 C; 231 G; 386 T; 0 U; 0 Other;

Query Match 99.7%; Score 1190.8; DB 3; Length 1194;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1192; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCATGAGATATAAATAAATCAGTGTGTTAATTTTCAGGTGACCAAGACTTAGAAGCAAAAC 60
Db 1 TCATGAGATATAAATAAATCAGTGTGTTAATTTTCAGGTGACCAAGACTTAGAAGCAAAAC 60

Qy 61 CTTATAGATCATGAGGTGGTGAATCATCATCATCAGTGTGCTCCAAACCGAGAAC 120
Db 61 CTTATAGATCATGAGGTGGTGAATCATCATCATCAGTGTGCTCCAAACCGAGAAC 120

Qy 121 TATAAAGGTGGCTTCGTGCTCCATATACGTAATCTTTGTCCTTTTGGCCAGCACTA 180
Db 121 TATAAAGGTGGCTTCGTGCTCCATATACGTAATCTTTGTCCTTTTGGCCAGCACTA 180

Qy 181 GCTACAAATCTGGGTAGATTCTATGAAATGGAGAAATAGCACATATGTTGTAACA 240
Db 181 GCTACAAATCTGGGTAGATTCTATGAAATGGAGAAATAGCACATATGTTGTAACA 240

Qy 241 CTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTGTTCCGCTTCTTTTCGAAATCAGG 300
Db 241 CTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTGTTCCGCTTCTTTTCGAAATCAGG 300

Qy 301 CAACCCAAATCAACAGATACAAATTCAGTCAGTCCCTTCCTTCACCAACCTTTGCATCG 360
Db 301 CAACCCAAATCAACAGATACAAATTCAGTCAGTCCCTTCCTTCACCAACCTTTGCATCG 360

Qy 361 GTTTACTTGTGACTGAGTCTAGTGTCCGCTTATGTTTGTCTGAGTAGGGTTG 420
Db 361 GTTTACTTGTGACTGAGTCTAGTGTCCGCTTATGTTTGTCTGAGTAGGGTTG 420

Qy 421 CTCTACTTACAGTCTCTACTTTTCTCCTCATCTGTTGGCTCACAGTTGGCCTTCACTGCC 480
Db 421 CTCTACTTACAGTCTCTACTTTTCTCCTCATCTGTTGGCTCACAGTTGGCCTTCACTGCC 480

Qy 481 TTTTCTCATATTTCTTAACTCGAAAGTTCACCTCTTTGATAGTCACTTTCTGCTT 540
Db 481 TTTTCTCATATTTCTTAACTCGAAAGTTCACCTCTTTGATAGTCACTTTCTGCTT 540

Qy 541 CTCTCTACTGTATCTCTGCTCTTCTGTTGCTCAACACTGATTCAGAAACTCAACTAAT 600
Db 541 CTCTCTACTGTATCTCTGCTCTTCTGTTGCTCAACACTGATTCAGAAACTCAACTAAT 600

Qy 601 GTATCTAGACAGTATGTGATCGGGTTTCAATGTACCATCGGTGCTTCGCTGGGATT 660
Db 601 GTATCTAGACAGTATGTGATCGGGTTTCAATGTACCATCGGTGCTTCGCTGGGATT 660

Qy 661 GGACTGTTACTATCTCTGATCAAAATGCTCTTCAGAAAGTTTTCAGAAACATACATCC 720
Db 661 GGACTGTTACTATCTCTGATCAAAATGCTCTTCAGAAAGTTTTCAGAAACATACATCC 720

Qy 721 TCAGCAGTTCAGGACTTGGCCATTTACAGTCTCTAGTTCGAGTGTGTAGTTCCTATA 780
Db 721 TCAGCAGTTCAGGACTTGGCCATTTACAGTCTCTAGTTCGAGTGTGTAGTTCCTATA 780

Qy 781 GGACTTTTTCAGTGGAGAGTGGAAACTTTTGGCCAAAGTATGAGATACTAACAATC 840
Db 781 GGACTTTTTCAGTGGAGAGTGGAAACTTTTGGCCAAAGTATGAGATACTAACAATC 840

Qy 841 GGGAAAGTGTATATGTTTTCAGTCTGAGTTCGCGAGCTATTTCTGCGAGTCTACACT 900
Db 841 GGGAAAGTGTATATGTTTTCAGTCTGAGTTCGCGAGCTATTTCTGCGAGTCTACACT 900

Qy 901 CCTGCTTGTGGGATTCATCTTCAGGATCATCTCTGTGTTCTCCAAATTCATCAACAGCT 960
Db 901 CCTGCTTGTGGGATTCATCTTCAGGATCATCTCTGTGTTCTCCAAATTCATCAACAGCT 960

Db 901 CTTGGTCTCTGGGATTGATCTTCAGTCACTCTCTGTGTTCTCCAATTCATAACAGCT 960

Qy 961 GTGGGATTGCTTAGTTCAGTTCGGGAGTAGTATTTTCCATGATAGAAATGGACGCA 1020
Db 961 GTGGGATTGCTTAGTTCAGTTCGGGAGTAGTATTTTCCATGATAGAAATGGACGCA 1020

Qy 1021 TCCAAAATCTTCTCCATTATTTTAGCTATCTGGGGCTTCTTTTCATTCTGTCATCAGCAC 1080
Db 1021 TCCAAAATCTTCTCCATTATTTTAGCTATCTGGGGCTTCTTTTCATTCTGTCATCAGCAC 1080

Qy 1081 TACCTCGACGAAAGAGTTGAATACCTAGGCACACACAGTCTGTAGGAGATCTTCACTTA 1140
Db 1081 TACCTCGACGAAAGAGTTGAATACCTAGGCACACACAGTCTGTAGGAGATCTTCACTTA 1140

Qy 1141 CTTGTTGAGGAAGTCCACAAACATACAAAGTGTGTGATCAAAAGCATATTTCC 1194
Db 1141 CTTGTTGAGGAAGTCCACAAACATACAAAGTGTGTGATCAAAAGCATATTTCC 1194

RESULT 2
ABZ14407
ID ABZ14407 standard; DNA; 3387 BP.

XX ABZ14407;
XX AC
XX ABZ14407;
XX AC
XX 21-JAN-2003 (first entry)
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2212.
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX OS Arabidopsis thaliana.
XX PN W0200216655-A2.
XX PD 28-FEB-2002.
XX PF 24-AUG-2001; 2001WO-US026685.
XX PR 24-AUG-2000; 2000US-0227866P.
XX PR 26-JAN-2001; 2001US-0264647P.
XX PR 22-JUN-2001; 2001US-0300111P.
XX PA (SCRI) SCRIPPS RES INST.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PI Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
XX PT Identifying a stress condition to which a plant cell has been exposed and
XX PS producing plants with increased tolerance to these abiotic stresses.
XX Claim 144; SEQ ID NO 2212; 577pp + Sequence Listing; English.
XX The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising: (a) contacting nucleic acid
XX representative of expressed polynucleotides in the plant cell with an
XX array or probes representative of the plant cell genome; and (b)
XX detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX in methods of the invention. Note: The sequence data for this patent is
XX not represented in the printed specification but is based on sequence
XX information supplied to Derwent by the European Patent Office
XX
SQ Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;

Query Match 85.1%; Score 1015.8; DB 6; Length 3387;
Best Local Similarity 95.1%; Pred. No. 1.1e-300;
Matches 1081; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

QY 38 GTGACCGAACTTAGAAGCAAAACCTTATAGATCATGAGGTGGTAACTGAATCATCATCAT 97
Db 1205 GTGACCGAACTTAGAAGCAAAACCTTATAGATCATGAGGTGGTAACTGAATCATCATCAT 1264

QY 98 CAGCTGTGCTCAAAACCGAGAACTATATAAGGTGGCTTGGTGTCTCCATATAGTAACTCT 157
Db 1265 CAGCTGTGCTCAAAACCGAGAACTATATAAGGTGGCTTGGTGTCTCCATATAGTAACTCT 1324

QY 158 TTGTCTCTTTTGGCCAGCCACTAGCTACAAATCTGGGTAGATTGTACTATGAAAAATGGAG 217
Db 1325 TTGTCTCTTTTGGCCAGCCACTAGCTACAAATCTGGGTAGATTGTACTATGAAAAATGGAG 1384

QY 218 GAAATAGCAGATATGTGGTAAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTGTCTGT 277
Db 1385 GAAATAGCAGATATGTGGTAAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTGTCTGT 1444

QY 278 TCCGCTCTTTTCTCGAATCAGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCC 337
Db 1445 TCCGCTCTTTTCTCGAATCAGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCC 1504

QY 338 CTTCTCTTCCACCCCTTGATCGGTTTACTTTGGCACTGGACTGTAGTGTCCGCTTATG 397
Db 1505 CTTCTCTTCCACCCCTTGATCGGTTTACTTTGGCACTGGACTGTAGTGTCCGCTTATG 1564

QY 398 CTTATTGTCTGCAGTAGGGTTGCTTACTTACCAGTCTCTACTTTCTCCCTCATCTTGG 457
Db 1565 CTTATTGTCTGCA----- 1578

QY 458 CCTCAGATGGGCTTCACTGCTTTTCTCATATTTCTTAATCTGCAAAAGTTCACTC 517
Db 1579 -----TTGGGCTTCACTGCTTTTCTCATATTTCTTAATCTGCAAAAGTTCACTC 1630

QY 518 CTTTGATAGTCAGTCTTTGCTCTCTCACTGTATCTCTGTCTTCTTGTGGTCAACA 577
Db 1631 CTTTGATAGTCAGTCTTTGCTCTCTCACTGTATCTCTGTCTTCTTGTGGTCAACA 1690

QY 578 CTGATTTCAGAAAACTCAACTAATGTATCTAGAGTACAGATGTGATCGGGTTTCATATGA 637
Db 1691 CTGATTTCAGAAAACTCAACTAATGTATCTAGAGTACAGATGTGATCGGGTTTCATATGA 1750

QY 638 CCATCGGTGCTCGCTGGATGGACTGTACTATCTGTATCAAAATGCTCTTCAGAA 697
Db 1751 CCATCGGTGCTCGCTGGATGGACTGTACTATCTGTATCAAAATGCTCTTCAGAA 1810

QY 698 AAGTTTTCAGAAAGCATACATCTCAGCAGTCAGGACTTGGCCATTTACCGAGTCTCTAG 757
Db 1811 AAGTTTTCAGAAAGCATACATCTCAGCAGTCAGGACTTGGCCATTTACCGAGTCTCTAG 1870

QY 758 TTGCGAGTTGTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGTGGGAACTTTTGCCAA 817
Db 1871 TTGCGAGTTGTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGTGGGAACTTTTGCCAA 1930

QY 818 GTGAGATGAGAAACTCAAACTCGGAAAGTGTATATGTTTGACTTTAGCTCGGAG 877
Db 1931 GTGAGATGAGAAACTCAAACTCGGAAAGTGTATATGTTTGACTTTAGCTCGGAG 1990

QY 878 CTATTTTCTGGCAAGTCTACACTCTGCTGTGGGATTGATCTTCGAGTCACTCTCG 937
Db 1991 CTATTTTCTGGCAAGTCTACACTCTGCTGTGGGATTGATCTTCGAGTCACTCTCG 2050

QY 938 TGTCTTCCAATTCCATAACAGCTGTGGGATTTGGCTATAGTTCAGTTCGGGAGTGATAG 997
Db 2051 TGTCTTCCAATTCCATAACAGCTGTGGGATTTGGCTATAGTTCAGTTCGGGAGTGATAG 2110

QY 998 TTTTCCATGATGAATGGAGCATCCAAAATCTTCTCCATTTATTTAGCTATCTCGGGCT 1057
Db 2111 TTTTCCATGATGAATGGAGCATCCAAAATCTTCTCCATTTATTTAGCTATCTCGGGCT 2170

QY 1058 TCCTTTTCATTTCGCTATCAGCACTACCTCGAGAAAGAGTTGAATATAGCCACAAA 1117
Db 2171 TCCTTTTCATTTCGCTATCAGCACTACCTCGAGAAAGAGTTGAATATAGCCACAAA 2230

QY 1118 GTGCTGTAGGAGATCTTCACTTACCTGTGTGAGGAAGGTCAACAACATACAAAGTG 1174
Db 2231 GTGCTGTAGGAGATCTTCACTTACCTGTGTGAGGAAGGTCAACAACATACAAAGTG 2287

RESULT 3
ADA68246
ID ADA68246 standard; DNA; 3387 BP.
XX
AC ADA68246;
XX
DT 20-NOV-2003 (first entry)
XX
DE Arabidopsis thaliana gene, SEQ ID 908.
XX
KW Plant; bacterial infection; fungal infection; viral infection; da.
XX
OS Arabidopsis thaliana.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 6; SEQ ID NO 908; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;

Query Match 85.1%; Score 1015.8; DB 8; Length 3387;
Best Local Similarity 95.1%; Pred. No. 1.1e-300;
Matches 1081; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

QY 38 GTGACCGAACTTAGAAGCAAAACCTTATAGATCATGAGGTGGTAACTGAATCATCATCAT 97
Db 1205 GTGACCGAACTTAGAAGCAAAACCTTATAGATCATGAGGTGGTAACTGAATCATCATCAT 1264

QY 98 CAGCTGTGCTCAAAACCGAGAACTATATAAGGTGGCTTGGTGTCTCCATATAGTAACTCT 157
Db 1265 CAGCTGTGCTCAAAACCGAGAACTATATAAGGTGGCTTGGTGTCTCCATATAGTAACTCT 1324

QY 158 TTGTCTCTTTTGGCCAGCCACTAGCTACAAATCTGGGTAGATTGTACTATGAAAAATGGAG 217
Db 1325 TTGTCTCTTTTGGCCAGCCACTAGCTACAAATCTGGGTAGATTGTACTATGAAAAATGGAG 1384

QY 218 GAAATAGCAGATATGTGGTAAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTGTCTGT 277
Db 1385 GAAATAGCAGATATGTGGTAAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTGTCTGT 1444

```
KW 278 TCCGCTTCTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCCC 337
XX 1445 TCCGCTTCTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCCC 1504
OS 338 CTTTCTTCAACACCCCTTCGATCGGTTTACTTGTGCACTGGAGTGTCTAGTGTCCGCTTATG 397
XX 1505 CTTTCTTCAACACCCCTTCGATCGGTTTACTTGTGCACTGGAGTGTCTAGTGTCCGCTTATG 1564
PD 398 CTTATTTCTGCGAGTAGGGTGTCTACTTACCAGTCTCTACTTTCTCCCTCACTCTGG 457
XX 1565 CTTATTTCTGCGA----- 1578
XX 458 CCTCACACTTGGCCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTC 517
XX 1579 -----TTGGCCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTC 1630
XX 518 CTTTGATAGTCAGTCTTTTGTCTTCCTCACTGATCTCTGCTCTTCTTGCTCAACA 577
XX 1631 CTTTGATAGTCAGTCTTTTGTCTTCCTCACTGATCTCTGCTCTTCTTGCTCAACA 1690
XX 578 CTGATTCAGAAAACCTCAACTAATGATCTAGAGTACAGTATGTGATCGGGTTCAATGTA 637
XX 1691 CTGATTCAGAAAACCTCAACTAATGATCTAGAGTACAGTATGTGATCGGGTTCAATGTA 1750
XX 638 CCATCGGTGCTTCGCTGGATTGGACTGTTACTATCTCTGATACAAATGCTCTCAGGA 697
XX 1751 CCATCGGTGCTTCGCTGGATTGGACTGTTACTATCTCTGATACAAATGCTCTCAGGA 1810
XX 698 AAGTTTTTCAGGAAGATACATCCTCAGCAGTCAAGGACTTGGCCATTACCAGTCTCTAG 757
XX 1811 AAGTTTTTCAGGAAGATACATCCTCAGCAGTCAAGGACTTGGCCATTACCAGTCTCTAG 1870
XX 758 TTGCGAGTTGTAGTCTCTAGAGACTTTTTCGAAGTGGAGAGTGGGAAACTTTTGCAA 817
XX 1871 TTGCGAGTTGTAGTCTCTAGAGACTTTTTCGAAGTGGAGAGTGGGAAACTTTTGCAA 1930
XX 818 GTGAGATGAGAACTACAACTCGGAAAGTGTATATGTTTTCAGTTAGCTTCGGCAG 877
XX 1931 GTGAGATGAGAACTACAACTCGGAAAGTGTATATGTTTTCAGTTAGCTTCGGCAG 1990
XX 878 CTATTTCTCGGAAGTCTACACTCCTGCTGTTGGGATTGATCTTCAGTCACTCCTCTG 937
XX 1991 CTATTTCTCGGAAGTCTACACTCCTGCTGTTGGGATTGATCTTCAGTCACTCCTCTG 2050
XX 938 TGTCTTCCAAATTCATACAGCTGTGGATTGCCATATAGTTCAGTTCGGCAGTGTATAG 997
XX 2051 TGTCTTCCAAATTCATACAGCTGTGGATTGCCATATAGTTCAGTTCGGCAGTGTATAG 2110
XX 998 TTTTCCATGATAGATGAGCGCATCCAAATCTCTCCATTTATTTTAGCTATCTCGCGCT 1057
XX 2111 TTTTCCATGATAGATGAGCGCATCCAAATCTCTCCATTTATTTTAGCTATCTCGCGCT 2170
XX 1058 TCCTTTCAATTCGCTATCAGCACTACCTCGAGAAAAGAGTTGAATACTAGCCACACAA 1117
XX 2171 TCCTTTCAATTCGCTATCAGCACTACCTCGAGAAAAGAGTTGAATACTAGCCACAA 2230
XX 1118 GTGCTGTAGGAGATCTTCACTACCTGTGTAGGAGAGTGCACAAACATACAAAGTG 1174
XX 2231 GTGCTGTAGGAGATCTTCACTACCTGTGTAGGAGAGTGCACAAACATACAAAGTG 2287
```

RESULT 4

AAA97922

ID AAA97922 standard; DNA; 1293 BP.

XX

AC AAA97922;

XX

DT 19-JAN-2001 (first entry)

XX

DE A. thaliana PUP1 DNA #4.

XX

KW PUP1; transgenic plant; nucleobase transporter; apical dominance;

flowering behaviour; senescence; pesticide distribution; ds.

Arabidopsis thaliana.

DE19907209-A1.

24-AUG-2000.

19-FEB-1999; 99DE-01007209.

19-FEB-1999; 99DE-01007209.

(FROM/) FROMMER W.

Gillissen B, Buerkle L, Andre B, Frommer WB;

WPI; 2000-566202/53.

Nucleic acid, useful for producing transgenic plants with altered nucleobase transport, encodes a nucleobase transporter protein of Arabidopsis thaliana.

Claim 1f; Page 13; 24pp; German.

This invention describes a novel nucleic acid encoding a plant nucleobase transporter (I). (I) is produced by complementation of a nucleobase transporter (NBT)-defective host cell with a plant gene bank by selection of NBT-positive cells. (I) is used to isolate homologous sequences from bacteria, fungi, plants, animals and humans, for expression of the encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting expression of (II) (when in antisense orientation), and to produce transgenic crop plants. The transgenic plants have modified nucleobase transport properties, e.g. altered affinity and substrate specificity that may result in more efficient nucleobase transport in leaves, changes in apical dominance, flowering behaviour and senescence, or improved distribution of pesticides. This sequence encodes the Arabidopsis thaliana PUP1 protein which is described in the method of the invention

SQ Sequence 1293 BP; 345 A; 288 C; 247 G; 413 T; 0 U; 0 Other;

Query Match

68.1%; Score 813.4; DB 3; Length 1293;

Best Local Similarity 81.6%; Pred. No. 9.6e-239;

Matches 969; Conservative 0; Mismatches 206; Indels 12; Gaps 2;

Qy 20 TGAGTGTAAATTTTCAGGTGACCGAGCACTAGAGCAAACTATAGATCATGAGTGG 79

Db 107 TGTTTTTCATGTTTTCAGATCATAAACAATAGAGCAAACTTCAAGGTCAGGAGAAA 166

Qy 80 TAACT-----GAATCATCATCATCAGCTGCCTCAAAACGAGAACTATAAAAGGT 130

Db 167 TGAATACCACTGGAATCGAATCTTCGTCGCTACCTCAATCGAAGAACTATAAGAAAT 226

Qy 131 GGCTTCGTGTCTCATATACGTAAATCTTTTCCTCTTTTCCAGCCACTAGTCAAAATTC 190

Db 227 GGCTTCGTAAATTCATTTACGTGTCTTCGTCTTGTCTGCTGCTGCAAGCACTTTCTACAATTT 286

Qy 191 TGGTAGATTTGTACTATGAAAATGGAGAAATAGCAGATATGTTGAACACTTCTTCAAC 250

Db 287 TGGCAGAGTTTACTATGAAAATGGTGGGAAGAGTACATGATGGGGAACACTTGTCCAAC 346

Qy 251 TCATTGGCTTCCCTGTACTGTTCTGTTCCGCTTCTTTTCTCGAATCAGGCAACCCAAAT 310

Db 347 TAATCGGCTTCCCTGTCTGTTCTCTTCTGCTTCTTTTCCCAACCAAAAATCCCAAC 406

Qy 311 CAACAGATACAAATTTTCAGTCAAGTCCCTTCCCTTCAACCCCTTGCATTCGGTTTACTTGT 370

Db 407 CAACAGAGCAGATTTTCAGAAAAGTCTCTTCCCTTCAACATTTCTTGGATCAGTTTACATCG 466

Qy 371 GCATGCACTGCTAGTGTCCGCTTATGCTTATTTGCTGAGTAGGTGCTCTACTTAC 430

Db 467 TTACTGCACTATTAGTGTCTGCTAAACTCTATATGCTCTGTTGGTTTACTACTTAC 526

Qy 431 CAGTCTCTACTTTCTCCCTCATCTTGGCCTCACAGTTGGCCTTCACTGCTTTTCTCAT 490

Db 527 CAGTTTCTACTTTCTCCCTCATCTTGGCCCTCACAATTGGCTTCACTGCTTCTTCTCAT 586
Qy 491 ATTTCCTTAACTCGCAAAAGTTCACTCCCTTTGATAGTCAGTTCTTTTCTTCTCTCACTG 550
Db 587 ATTTTCTAACTCGCAGAAGTTCAACCTTTCAATGTGAATCTCTGTTTCTCTTACTA 646
Qy 551 TATCCTCTGCTCTTCTTGTGGTCAACACTGATTCAGAAAACCTCAATAATGATCTAGAG 610
Db 647 TTTCTCTGCGCTCTCTGCTGCTCAACACTGATTCGGAACAACAGCAAAAGTGTCTAGAG 706
Qy 611 TACAGTATGATCGGGTTCAATGATACCATCGGTCTTCCGCTGGGATGGACTGTAC 670
Db 707 TAAATATGATAGGATAATATGATACCATGCTGCTGCTGGGATTTGATTTGCTGC 766
Qy 671 TATCTCTGATACAAATGCTCTTCCAGGAAGTTTTCAGGAAGCATATCATCTCAGCAGTCA 730
Db 767 TATCCCTGGTACAACTGATCTCAGGAGGTTTAAAGAGCAAACTCTCAACGGTCA 826
Qy 731 CGGACTTGGCCATTTACCAAGTCTCTAGTTCCGAGTTGTGTAGTTCTCATAGGACTTTTGG 790
Db 827 CTGACTTGGTGGCTTACCAATCTCTAGTTGCAAGCTGTGTGTTCTCATAGGACTTTTGG 886
Qy 791 CRAAGTCGAGTGGGAACTTTCCCAAGTGAGATGAGAACTACAAACTCGGAAAGTGT 850
Db 887 CAAGCGGGAGTGGAAAACCTTTAACAAGTGAGATGGAAAACCTACAAACTGGGGAAAGTGC 946
Qy 851 CATATGTTTGACTTTAGCTTGGCTGGCAGCTATTTCTTGGCAAGTCTACACTCTCTGGTCTTG 910
Db 947 CATACGTTATGACTTTGGCTCGATAGCTATTTCTTGGCAAGTCTACACCATTTGGCGTGC 1006
Qy 911 TGGGATGATCTTCGAGTCACTCTCTGTGTTCTTCCAAATCCATAACAGCTGTGGGATTCG 970
Db 1007 TGGGACTGATCTTTGAGTCACTCTCTGTGTTCTTCCAAATCCATAACTGCTGTGGGATTCG 1066
Qy 971 CTATAGTTCCAGTTGGGCAAGTATGATTTTCCATGATAGATGAGCGCATCAAAATCT 1030
Db 1067 CTATAGTTCCAGTTGAGCAGTGAATGTTTCCATGATAGATGAGCGGTCAAAAGATCT 1126
Qy 1031 TCTCCATATTTTAGCTATCTGGGCTTCTTCTTCAATGCTCTATCAGCACTACCTCGAGC 1090
Db 1127 TCTCCATATTTTAGCTATCTGGGATCAATTTCAATTTGCTATCAGCACTACCTCGAGC 1186
Qy 1091 AAAAGAAGTTGAATACTAGCCACACAAAGTGTGTAGGAGATCTTCAAT---CTACCTGTTG 1147
Db 1187 AAAAGAAGTTGAAGTACGACACCAAGTCTGTAGGAGATCTTCAATCTACTACTGCTG 1246
Qy 1148 ACGAAGTCAACAACATACAAAGTGTGTGATCAAGCATATTTCC 1194
Db 1247 AGGAAGTCAACAACATACATAGTGTATGATCAAAACATATTTCC 1293

RESULT 5

ABX56692
ID ABX56692 standard; DNA; 592 BP.

XX AC ABX56692;

XX AC 20-FEB-2003 (first entry)

XX DE Arabidopsis thaliana polynucleotide #44.

XX KW Thale cress; gene; ds; genetic manipulation; plant; biosynthesis;
XX KW genetic modification; environmental stress; disease resistance;
XX KW fungicide; insecticide; stress tolerance.

XX OS Arabidopsis thaliana.

XX PN US2002040489-A1.

XX PD 04-APR-2002.

XX PF 26-JAN-2001; 2001US-00770152.

XX 27-JAN-2000; 2000US-0178503P.
XX (GORL/) GORLACH J.
PA (ANY/) AN Y.
PA (HAMI/) HAMILTON C M.
PA (PRIC/) PRICE J L.
PA (RAIN/) RAINES T M.
PA (YUY/) YU Y.
PA (RAME/) RAMEKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHAW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRIC/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX Goriach J, An Y, Hamilton CW, Price JL, Raines TM, Yu Y;
PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX WPI; 2003-110410/10.
XX Novel Arabidopsis thaliana nucleic acid useful for identifying homologous
PT or related genes, and to create genetically modified and transgenic
PT organisms, such as plant cells and plants.
XX Claim 1; SEQ ID NO 44; 45pp; English.
XX The invention relates to Arabidopsis thaliana nucleic acid sequences. The
CC DNA sequences and the polypeptides they encode are useful for identifying
CC homologous or related genes, for producing compositions that modulate the
CC expression or function of the polypeptides, for mapping functional
CC regions of the protein, in diagnosis, for studying associated
CC physiological pathways, for genetic manipulation of cells, preferably
CC plant cells, in screening assays of various plant strains to determine
CC the strains that are capable of withstanding a particular disease or
CC environmental stress, for enhancing or inhibiting production of
CC biosynthetic products in plants and to create genetically modified and
CC transgenic organisms, such as plant cells and plants. Transgenic plants
CC are useful for introducing or improving disease resistance and stress
CC tolerance in plants, screening biologically active agents, such as
CC fungicides and insecticides, and for elucidating biochemical pathways.
CC Sequences ABX56649-ABX57647 represent Arabidopsis thaliana
CC polynucleotides of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html
XX SQ Sequence 592 BP; 135 A; 158 C; 96 G; 203 T; 0 U; 0 Other;

Query Match 38.5%; Score 459.8; DB 10; Length 592;
Best Local Similarity 91.6%; Pred. No. 2.6e-130;
Matches 501; Conservative 0; Mismatches 37; Indels 9; Gaps 1;

Qy 37 GGTGACCAAGAACTTAGAAGCAAACTTATAGATCATGAGGTGGTAACCTGAATCATCATCA 96
Db 55 GGTGACCAAGAACTTAGAAGCAAACTTCTAGATCATGAG-----GAACTGAATCA 105
Qy 97 TCAGCTGTGCTCAAAACGAGAACTATATAAGGTGGTTCGTCTTCCATATAGTAATC 156
Db 106 TTTTCAGTACCTCAAAACGAGAACTGTAAAGGTGGCTCCGCTGCTCCATATACGCAATC 165
Qy 157 TTTTCTCTCTTTTGGCAGCCACTAGTACAAATTTCTGGGTAGATTTGTACTATGAAATGGA 216
Db 166 TTTGTGTCATCTTCTGCCAACCACTTGTCTACAGTTCTGGGTAGACTGTACTATGAAATGGA 225

QY 217 GGAAATAGACATATGTGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTG 276
|||
Db 226 GGGAAAGACATATGTGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTG 285
|||
QY 277 TTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAAATTTTCAGTCAGTCC 336
|||
Db 286 TTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAAATTTTCAGTCAGTCC 345
|||
QY 337 CCTTCCTTACACACCTTGCATTCGGTTTACTTGTGCACTGGACGTAGTGTCCGCTTAT 396
|||
Db 346 CCTTCCTTACACACCTTGCATTCGGTTTACTTGTGCACTGGACGTAGTGTCCGCTTAT 405
|||
QY 397 GCTTATTTGTCTGCAGTAGGGTGTCTCTACTTACCAAGTCTCTACTTTCCTCCTCATCTTG 456
|||
Db 406 GCTTATTTGTCTGCAGTTGGGTGCTTTACTTACCAAGTCTCTACTTTCCTCCTCATCTTG 465
|||
QY 457 GCCTCACAGTTGGCCTTCACTGCCTTTTCTCATATTTTCTTAACTCGCAAAAGTTCAC 516
|||
Db 466 GCCTCACAGTTGGCCTTCACTGCCTTTTCTCATATTTTCTTAACTCGCAAAAGTTCAC 525
|||
QY 517 CCTTGTAGTACAGTCTTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 576
|||
Db 526 CCTTGTAGTACAAATCTTTGTTTCTCTTACGGTTTCTCTGCTCTCTCTCTCTCTCTCT 585
|||
QY 577 ACTGATT 583
|||
Db 586 ACTGATT 592
|||

RESULT 6

AAA97925

ID AAA97925 standard; DNA; 1071 BP.

XX

AC AAA97925;

XX

19-JAN-2001 (first entry)

XX

A. thaliana PUP1 DNA #7.

XX

PUP1; transgenic plant; nucleobase transporter; apical dominance;

XX flowering behaviour; senescence; pesticide distribution; ds.

XX

OS Arabidopsis thaliana.

XX

FN DE19907209-A1.

XX

PD 24-AUG-2000.

XX

PF 19-FEB-1999; 99DE-01007209.

XX

PR 19-FEB-1999; 99DE-01007209.

XX

PA (FROM/) FROMMER W.

XX

PI Gillissen B, Buerkle L, Andre B, Frommer WB;

XX

DR WPI; 2000-566202/53.

XX

Nucleic acid, useful for producing transgenic plants with altered

XX nucleobase transport, encodes a nucleobase transporter protein of

PT Arabidopsis thaliana.

XX

PS Claim 2; Page 15; 24pp; German.

XX

This invention describes a novel nucleic acid encoding a plant nucleobase

CC transporter (I). (I) is produced by complementation of a nucleobase

CC transport (NBT)-defective host cell with a plant gene bank by selection

CC of NBT-positive cells. (II) is used to isolate homologous sequences from

CC bacteria, fungi, plants, animals and humans, for expression of the

CC encoded protein (ii) in prokaryotic or eukaryotic cells, for inhibiting

CC expression of (ii) (when in antisense orientation), and to produce

CC transgenic crop plants. The transgenic plants have modified nucleobase

CC transport properties, e.g. altered affinity and substrate specificity

CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX
SQ Sequence 1071 BP; 307 A; 236 C; 208 G; 320 T; 0 U; 0 Other;

Query Match 34.0%; Score 405.8; DB 3; Length 1071;
Best Local Similarity 64.5%; Pred. No. 1.4e-113;
Matches 624; Conservative 0; Mismatches 337; Indels 6; Gaps 1;

QY 164 TCTTTTGGCAGCCACTAGCTACAAATCTGGGTAGATTGTACTATGAAAATGGAGAAATA 223
|||
Db 47 TCATTGGCCAAATCAGTTGCTACAAATCTGGGACACTATACTATGAAAATGGAGAAACA 106
|||
QY 224 GCACATATGTGGAACACTTCTTCAACTCAATGGCTTCCCTGTACTGGTTCCTGTCGCT 283
|||
Db 107 GCAATGGCTAGCAACGGTAGTTTTCAGCTTGTAGCTTTCCTATTTCTACTTCCATATCATC 166
|||
QY 284 TCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCCTTCCCT 343
|||
Db 167 TCTTGTGAGTCA-----AAACACATACAACTCAGAGAGATGGCAAAATTAACCTCAC 220
|||
QY 344 TCACCACTTGCATCGGTTTACTTGTGCACTGGAAGTGTAGTGTGCGCTTATGCTTATT 403
|||
Db 221 TTAGGAACCGTGCATTAGTTTACATAGTGTGCTTGCAGCTTCTTGTAGGAGCAGCTTGCTACC 280
|||
QY 404 TGCTGCAGTAGGGTGTCTCTATACAGTCTCTACTTTTCTCCCTCATCTTGGCTCAC 463
|||
Db 281 TATATTCATTTGGACTGCTTTAGCTACCTGCTTTTCTACCTTTCCTGATCTGTGCATCAC 340
|||
QY 464 AGTTGGCTTCACTGGCTTTTCTCATATTTTCTTAATCTGCAAAAGTTCACCTCTTTGA 523
|||
Db 341 AGTTAGCTTTCACGCTTCTTCTTATTTTACTCAACTCACAATAACTTACTTCTCTATCA 400
|||
QY 524 TAGTCAGTCTTCTGCTTCTCCTCACTGTATCTCTGCTCTGCTTCTTGTGGTCAACACTGATT 583
|||
Db 401 TTTTGAATCTCTTTTCTCTCTCATATATCTTCCACCTCTTGCATTTTAATAACGAGG 460
|||
QY 584 CAGAAAACCTCAACTAATGTATCTAGAGTACAGTATGTGATCGGGTTCATATGATACCATCG 643
|||
Db 461 AATCAGATTCCAAAAAAGTTTACAAAAGGAGATGTCTCAAGGTTCGTATGCACTCCGTTG 520
|||
QY 644 GTGCTTCGCTGGGATTTGGACTGTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTT 703
|||
Db 521 GTGCATCTGCTGGGTTTGGTCTACTCTTATCCCTTACAACTAGCCTTTCGTAAAGTTT 580
|||
QY 704 TCACGAAGCATACATCCTCAGCAGTCACTGGCACTTTGGCCATTTTACCAGTCTCTAGTTGCGA 763
|||
Db 581 TAAAGAAGCAAACTTTTCTCAGAAAGTTATAAATATGATAATCTACATGAGTCTAGTGSCCA 640
|||
QY 764 GTTGTGTAGTTTCTCATAGGACTTTTTCGAAAGTGGAGAGTGGGAAACTTTGCCAAGTGAGA 823
|||
Db 641 GTTGTGTAGCTGGTGGGCTTTTGTCTAGTAGCGAGTGGAAACTTTTGAGCAGCTGAAA 700
|||
QY 824 TGAGAAACTACAACTCGGAAAGTGTCTATGTTTTTGTAGCTTTAGCCTCGGCACTGATT 883
|||
Db 701 TGGAAAACCTACAACTTTGGGAAGGTATCTTATGTATGAACCTTAGTGTGACAGCTGTTA 760
|||
QY 884 CCTGGCAAGTCTACACTCTTGGTCTTGTGGGATTTGATCTTTCGAGTCACTCTCTGTGTCT 943
|||
Db 761 CCTGGCAGGTATTTCTCCATCGGTTGCACAGACTGATCTTCGAGCTTTCTCCCTATTTCT 820
|||
QY 944 CCAATTCATACAGCTGTGGGATTTGCTATAGTTTCCAGTTGGGCACTGATAGTTTTC 1003
|||
Db 821 CAAATGCAATAAGCGCTTTTGGGACTTCCCGTGGTTCCTATCTCTGGCTGTCTATCATTTTCC 880
|||
QY 1004 ATGATAGAATGGAGCGCATCCAAAATCTTCTCCATTTTATGCTATCTCGGGCTTCCCTTT 1063
|||
Db 881 ATGACAAATGAACGGCTTAAAGGTGATTTCTATGATTTCTAGCTATTTTGGGGTTTCGTAT 940
|||
QY 1064 CATTTGCTTATCAGCACTACCTCGACGAAAGAAAGTTGAATATAGCCACACAAAGTCGTG 1123
|||

Db	941	CCTATGCTTACCACAAATATCTTGATGAACAACTTGAGAAAAGTAATGAAATACCAA	1007
Qy	1124	TAGGAGA	1130
Db	1001	CAACAGA	1007
RESULT 7			
ADG88190			
ID	ADG88190	standard; cDNA; 1047 BP.	
XX	AC	ADG88190;	
XX	DT	22-APR-2004 (first entry)	
XX	DE	A. thaliana RPP4-upregulated pathogen infection-related gene #632.	
XX	KW	Pathogen infection-related gene; plant; Peronospora parasitica; defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.	
XX	OS	Arabidopsis thaliana.	
XX	PN	WO200222675-A2.	
XX	PD	21-MAR-2002.	
XX	PF	14-SEP-2001; 2001WO-US028506.	
XX	PR	15-SEP-2000; 2000US-0232778P.	
XX	PR	22-JUN-2001; 2001US-0300183P.	
XX	PA	(SYGN) SYNGENTA PARTICIPATIONS AG.	
XX	PA	(UYNC-) UNIV NORTH CAROLINA.	
XX	PA	(GLAZ/) GLAZEBROOK J.	
XX	PA	(WANG/) WANG X. L.	
XX	PA	(DANG/) DANG J. L.	
XX	PA	(EULG/) EULGEM T.	
XX	PA	(ZHUT/) ZHU T.	
XX	PI	Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;	
XX	DR	WPI; 2002-292409/33.	
XX	PT	Novel isolated polynucleotide, useful for conveying pathogen resistance to plants, and for identifying plants infected with a pathogen.	
XX	PS	Claim 3; SEQ ID NO 632; 605pp; English.	
XX	CC	The invention relates to 691 Arabidopsis thaliana genes (ADG87559--ADG87557) whose expression is altered in response to pathogen infection, and to homologues of these genes from other plants or fungi, especially from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape), cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The expression of genes of the invention was upregulated or downregulated in Arabidopsis plants infected with the oomycete Peronospora parasitica, indicating that they play a role in defence mechanisms. The genes of the invention are regulated by RPP7 or RPP8 which act via unconventional signalling cascades, or by the RPP4-dependent pathway. The invention also relates to polypeptides encoded by the pathogen infection-related genes; promoter motifs from pathogen infection-related genes (ADG88243-ADG88327); expression cassettes, host cells and pathogen-resistant transgenic plants and their progeny comprising a polynucleotide of the invention; and a method of identifying a plant cell infected with a pathogen. The polynucleotide sequences and methods of the invention are useful for identifying plants infected with a pathogen, and for conferring resistance to pathogens such as oomycetes, fungi, bacteria, viruses, nematodes and insects (e.g., aphids). The present sequence represents an Arabidopsis thaliana gene whose expression is altered in response to Peronospora parasitica infection. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published pct sequences.	

PR 30-APR-1999;	99US-0132048P.	PR 21-JUL-1999;	99US-0145086P.
PR 30-APR-1999;	99US-0132407P.	PR 21-JUL-1999;	99US-0145088P.
PR 04-MAY-1999;	99US-0132484P.	PR 22-JUL-1999;	99US-0145085P.
PR 05-MAY-1999;	99US-0132485P.	PR 22-JUL-1999;	99US-0145087P.
PR 06-MAY-1999;	99US-0132486P.	PR 22-JUL-1999;	99US-0145089P.
PR 06-MAY-1999;	99US-0132487P.	PR 22-JUL-1999;	99US-0145192P.
PR 07-MAY-1999;	99US-0132863P.	PR 23-JUL-1999;	99US-0145145P.
PR 11-MAY-1999;	99US-0134256P.	PR 23-JUL-1999;	99US-0145218P.
PR 14-MAY-1999;	99US-0134218P.	PR 23-JUL-1999;	99US-0145224P.
PR 14-MAY-1999;	99US-0134219P.	PR 26-JUL-1999;	99US-0145276P.
PR 14-MAY-1999;	99US-0134221P.	PR 27-JUL-1999;	99US-0145913P.
PR 14-MAY-1999;	99US-0134370P.	PR 27-JUL-1999;	99US-0145918P.
PR 18-MAY-1999;	99US-0134768P.	PR 27-JUL-1999;	99US-0145919P.
PR 19-MAY-1999;	99US-0134941P.	PR 28-JUL-1999;	99US-0145951P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146386P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146388P.
PR 24-MAY-1999;	99US-0135629P.	PR 02-AUG-1999;	99US-0146389P.
PR 25-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147038P.
PR 27-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147204P.
PR 28-MAY-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147192P.
PR 03-JUN-1999;	99US-0137528P.	PR 05-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149910P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139819P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 13-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.

PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match			26.5%; Score 316.8; DB 3; Length 1292;
Best Local Similarity			57.1%; Pred. No. 3.2e-86;
Matches 604; Conservative 0; Mismatches 442; Indels 12; Gaps 1;			
QY	53	AAGCAACCTTATAGATCATAGGTGTAAGTAACTGAATCATCATCATCAGCTGTGCCTCAAA	112
Db	47	AAGAATCCATCTTCCAGTAAATGGCGAACCTGAAGGAAATTTTCAACAGAGAGAGAA	106
QY	113	CCGAGAACTATAAAGTGCCTTCGTCTCCATATACGTAATCTTTGTCCTTTTGGC	172
Db	107	GTCAAGTACTCTTGAGGTAAAGAGTCTCTCTATGTCACTCTCTCTTAGCTGGAG	166
QY	173	AGCCACTAGCTACAATTTGGGTAGATTTGACTATGAAATGAGGAAATAGCACATATG	232
Db	167	AGACAATAGCACTCTCTTAGGTAGACTTTACTACGAAAGGCGTAAAGCACATGGC	226
QY	233	TGGTAACATTTCTTCAACTCATTTGGCTTCCGTGTACTGGTTCGTTTCGGCTCTTTTTC	292
Db	227	TCGAAACCTTGGTTTCACTTGTAGGTTTTCCTTTAAACCTTCTCTCTATTATTACTAA	286
QY	293	GAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCCCTTCTTCCACACC	352
Db	287	AGCTGAGCGCTCAAGACTAAACCAATACCAAAAAAACTACTCTCTCTCTTGACAC	346
QY	353	TTGCATCGGTTTACTTGTGCACCTGACTGTCTAGTGTCCGCTTATGTCTTGTCTGCAG	412
Db	347	TATCTTTAGTGATATTTGACTTGGCTTGGTGTCTGACATGTTATTTGTACTCAT	406
QY	413	TAGGTTGTCTACTTACAGTCTCTACTTCTCCCTCACTTGTGGCTCAGAGTTGGCT	472
Db	407	TTGGGCTACTTTACCTTCTCTCACTTCTCACTTCTCTTGTATCTCTGGTCCCAATTGGCT	466
QY	473	TCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTCACTCTCTTTGATAGTCAGTT	532
Db	467	TTAACGCGTCTTCTCTTACTTCTTAACTCAAAAAATCAACCAATTTATCTCAAT	526
QY	533	CTTTGCTTCTCTCACTGTATCTCTGCTCTTCTTGTGTCAACACTGATTCAGAA- - -	588
Db	527	CACCTGTTCTTTAAACATATCTTCTACACTTCTTGTATCCAAATGACCAAGATCTC	586
QY	589	-----AAGTCAAGTATAGTATCTAGAGTACAGATGTGATCGGGTTTATATGACCA	640
Db	587	CCTCTTCTACTTCAAAAGTCCGAGCCAAAGTCCAAAGTATGTGATGATACATCTGGCGG	646
QY	641	TCGGTCTTCCGTTGGGATGGACTGTCTTACTATCTCTGATACAAATGCTCTTCAGGAAG	700
Db	647	TCGGTAGCTCAGCTGGTATTTCTCTGGTCTTCTTTTAAAGATTAACGGTTTCAAAAGA	706
QY	701	TTTTCAGGAAGATACATCTCTCAGCAGTCCGGCTTTGGCCATTTTACCACTCTCTAGTTG	760
Db	707	TTCTAAGAAATACAACTCAAGGCTATTTTAGACATGGCCACATATCTCTCTATGTTAG	766
QY	761	CGAGTTGTGTAGTTCTCATAGGACTTTTTTGAAGTGGAGAGTGGGAAACTTTGCCAAGTG	820

Db	767	CTACTTGTAGTGTGGTAGGACTTTTTTGAAGTGGTGGTGGAAAAAGCTGAGTACAG	826
QY	821	AGATGAGAACTACAACTCGGGAAGTGTCTATATGTTTGTACTTTTAGCTCGGACGTA	880
Db	827	AAATGGAAGAGTTTCAACTAGGGAAGGCTCATACATTTTGATATAAACATCGGTTCAACGA	886
QY	881	TTTCTCTGCAAGTCTACACTCTCTGTCTTTGTTGGGATTCATCTTCGAGTCATCTCTGTGT	940
Db	887	TATCATGGCAAGCTTTTGTGTTGGAAGTGTGTTGTTGATTATCGAAGTTTTCATCGCTTT	946
QY	941	TCTCCAAATTCATAACAGCTGTGGGATTCGCTATAGTTCCAGTTGGCGGAGTATAGTTT	1000
Db	947	TTTCCAATGTCTAAGCACTCTTTGTTTACCAGTTGTGCTCTTCTTGTCTGTCTTCT	1006
QY	1001	TCCATGATAGATGGAGCATCCAAAATCTTCTCCATTTATTTTAGCTATCTCGGCTTCC	1060
Db	1007	TCCGTGATGAGATGAGTGAATCAAGTTGGTGGCAATGTTTGGCCATCTGGGATTTG	1066
QY	1061	TTTCAATTCCTTATCAGCACTACCTCGACGAAAAAGAG	1098
Db	1067	TTTCTTATGTTATCAGCATTATGTCATATGATAGAAAG	1104
RESULT 11			
AAC37289			
ID	AAC37289 standard; DNA; 1295 BP.		
XX	AAC37289;		
AC	AAC37289;		
XX	17-OCT-2000 (first entry)		
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 16846.		
DE	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX	Arabidopsis thaliana.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
XX	06-SEP-2000.		
PD	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-00301439.		
PF	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	30-APR-1999; 99US-0132407P.		
PR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	06-MAY-1999; 99US-0132487P.		
PR	07-MAY-1999; 99US-0132863P.		
PR	11-MAY-1999; 99US-0134256P.		
PR	14-MAY-1999; 99US-0134218P.		
PR	14-MAY-1999; 99US-0134219P.		
PR	14-MAY-1999; 99US-0134221P.		
PR	14-MAY-1999; 99US-0134370P.		

PR 18-MAY-1999;	99US-0134768P.	PR 27-JUL-1999;	99US-0145919P.
PR 19-MAY-1999;	99US-0134941P.	PR 28-JUL-1999;	99US-0145951P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146386P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146388P.
PR 24-MAY-1999;	99US-0135629P.	PR 02-AUG-1999;	99US-0146389P.
PR 25-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147038P.
PR 27-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147204P.
PR 28-MAY-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147192P.
PR 03-JUN-1999;	99US-0137528P.	PR 05-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149358P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140821P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0146814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.

PR	26-OCT-1999;	99US-01611359P.	
PR	26-OCT-1999;	99US-01611360P.	
PR	26-OCT-1999;	99US-01611361P.	
PR	28-OCT-1999;	99US-01611920P.	
PR	28-OCT-1999;	99US-01611922P.	
PR	28-OCT-1999;	99US-01611933P.	
PR	28-OCT-1999;	99US-0162142P.	
Query Match			
Best Local Similarity 57.1%; Score 316.8; DB 3; Length 1295;			
Matches 604; Conservative 0; Mismatches 442; Indels 12; Gaps 1;			
QY	53	AAGCAACCTTATAGATCATGAGTGGTAACTGAATCATCATCATGCTGGCTCAAA	112
Db	48	AAGAACTCCATCTTCACGTAAATGGCGAACCTGAAGGGAAATTTTCAACAGAAGAGAA	107
QY	113	CCGAGAACTATAAAGTGGCTTCGTCTCCATATACGTAATCTTTTGCCTCTTTGGC	172
Db	108	GTCAAGTACTCTTGAGGTTAAGAGTGTCTCTATGTCACTCTCTCTTAGCTGGAG	167
QY	173	AGCCACTAGCTACAATCTGGTGTAGATGTACTATGAAATGGAGGAAATAGCACATATG	232
Db	168	AGCAATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAAAGCGGTAAAGCACATGGC	227
QY	233	TGTTAACACTTCTTCAACTATTTGGCTTCCGTACTGTGTTCTGTTCCGCTCTTTTCTC	292
Db	228	TCGAAACCTTGGTTCAGCTTGTAGGGTTTCCTTTAAACCCCTTCCTTCTATTACTTAA	287
QY	293	GAATCAGGCAACCCAACTCAACAGATACAAATTTCACTCAGTCCCTCTTCACCAACC	352
Db	288	AGCTGAGCGCTCCAGACTAAACCAATTAACCAAAAACTACTCTCTCTCTTGACAC	347
QY	353	TTGCATCGGTTTACTTGTGCACCTGGACTGCTAGTGTCCGCTTATGCTTATTTGTCTGCAG	412
Db	348	TATCTTTAGTGATATTGACTTGGCTTGTGTTCTGGACATGTATTTTGTACTCAT	407
QY	413	TAGGGTGTCTACTTACAGTCTCTACTTCTCCCTCATCTTGGCTCACAGTGGCT	472
Db	408	TTGGGTACTTTTACCTTCCGTCTCAACTTCTCTTTGATCTCTGGTGCCTAATGGCTT	467
QY	473	TCAGTCCCTTTTCTCATATTTCTTAATCTCGCAAGTTTCACTCCTTTGATAGTCAGTT	532
Db	468	TTAACCCGCTTCTCTTACTTCTTAATCTCAAAAAATCAACCAATTTTACTCAAT	527
QY	533	CTTTGCTTCTCTCACTGATCTCTGCTCTTCTTGTGCTCAACACTGATTCAGAA- - -	588
Db	528	CACCTGTTCTTTAAACCATATCTTCACTCTTGTATCCACATGACCAAGATCTC	587
QY	589	-----NACTCAACTAATGATCTAGAGTACAGTATGATCGGGTTTATATGTACCA	640
Db	588	CCTCTTCTACTTCAAAAGTCCGACGCCAAGTCCAAGTATGTGATGGATACATCTGCGGG	647
QY	641	TCGGTCTTCCGCTGGATGGACTGTATTAATCTCTGATACAAATGCTCTTCAGGAAG	700
Db	648	TCGGTAGCTCAGCTGGTTATTCTCTGGTCTTCTTTTAAACAGATTACGGGTTTCAAAAGA	707
QY	701	TTTTACGAAGCATACATCTCTCAGCAGTCAAGCTTTGGCCATTTACCACTCTAGTTG	760
Db	708	TTCTAAGAAATACACATTTCAAGGCTATTTTAGACATGGCCACATATCCGCTATGGTAG	767
QY	761	CGAGTTGTGTAGTTCTATAGCACTTTTTCGAAGTGGAGAGTGGGAACTTTGCCAAGTG	820
Db	768	CTACTGTGTAGTTGGTAGGACTTTTGGGAAGTGGTGGGTGGAAAGCTGAGTACAG	827
QY	821	AGATGAAACTACAACTCGGAAAGTGTATATTTTGAATTTAGCTCGGAGCTA	880
Db	828	AAATGGAAGAGTTTCACTAGGGAAGCTCTACATATTTTGATAAATCTCGGTTTCAACGA	887
QY	881	TTTCTCTGGAAGTCTACACTCTCTGCTTGTGGGATTTGATCTTCAGTCACTCTGTGT	940
Db	888	TATCATGGCAAGCTTGTGTGATGGAGTGTGGTTGGTTGATATCGAAGTTTCATCGCTT	947
QY	941	TCTCCAATTCATAACAGCTGTGGGATTCCTATAGTTCCAGTTGGGGCAGTGATAGTTT	1000

Db	948	TTTCCAATGTCTAATAGCACTCTTTGTTTACCAGTTGTGCTGTCTTGTCTGTCTCT	1007
QY	1001	TCCATGATAGAAATGGAGCATCCAAAATCTTCTCCATTTATTTAGCTATCTCGGCTCC	1060
Db	1008	TCCGTGATGAGATGAGTGGATCAAGTTGATTTGCAATGTTTTTGGCCATCTGGGATTTG	1067
QY	1061	TTTCAATTCGTCTATCAGCACTACCTCGACGAAAGAAAG	1098
Db	1068	TTTCTTATGTTTATCAGCATTATGTCAATGATAGAAAG	1105
RESULT 12			
AAC36845			
ID	AAC36845 standard; DNA; 492 BP.		
XX	AAC36845;		
XX	17-OCT-2000 (first entry)		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 15272.		
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
XX	promoter; termination sequence; ss.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
PD	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-00301439.		
XX	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	30-APR-1999; 99US-0132407P.		
PR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	06-MAY-1999; 99US-0132487P.		
PR	07-MAY-1999; 99US-0132863P.		
PR	11-MAY-1999; 99US-0134256P.		
PR	14-MAY-1999; 99US-0134218P.		
PR	14-MAY-1999; 99US-0134219P.		
PR	14-MAY-1999; 99US-0134221P.		
PR	18-MAY-1999; 99US-0134370P.		
PR	19-MAY-1999; 99US-0134768P.		
PR	19-MAY-1999; 99US-0134941P.		
PR	20-MAY-1999; 99US-0135124P.		
PR	21-MAY-1999; 99US-0135353P.		
PR	24-MAY-1999; 99US-0135629P.		
PR	25-MAY-1999; 99US-0136021P.		
PR	27-MAY-1999; 99US-0136392P.		
PR	28-MAY-1999; 99US-0136782P.		
PR	01-JUN-1999; 99US-0137222P.		
PR	03-JUN-1999; 99US-0137528P.		
PR	04-JUN-1999; 99US-0137502P.		
PR	07-JUN-1999; 99US-0137724P.		

PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139453P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149375P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151338P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159631P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 25.6%; Score 305.8; DB 3; Length 492;
Best Local Similarity 90.0%; Pred. No. 4.8e-83;
Matches 341; Conservative 1; Mismatches 28; Indels 9; Gaps 1;

QY 37 GGTGACCAGAACTTAGAGCAAAACCTTATAGATCATGAGGTGTAATCATCATCA 96
Db 123 GGTGACCAGAACTTAGAGCAAAACCTTCTAGATCATGAG-----GAAACTGAATCA 173
QY 97 TCAGCTGTGCTCTCAAAACCGAGAACTATAAAGGTGGCTTCGTGTCTCCATATACGTAATC 156
Db 174 TTTTCAGTACTCANAGCAAACTGTAGAGGTGGCTCGTGTCTCCATATACGCAATC 233
QY 157 TTTGTCTCTTTTCCAGGCCACTAGCTACAAATTCGGGTAGATTGTACTATGAAAAATGGA 216
Db 234 TTTGTCTATCTCTGCCAACCACTTGTACAGTTCTGGGTAGACTGTACTATGAAAAATGGA 293
QY 217 GGAAATAGACATATGTGTTAACTCTTCAACTCATTTGGCTTCCTGTACTGTTCTG 276
Db 294 GGGAAAAGCACATATGTGTAACACTTCTTCAACTCATTTGGCTTCCTGTACTGATCTC 353
QY 277 TTCGGCTCTCTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTCAGTCAAGTCC 336
Db 354 TTCGGCTCTCTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTCAGTCAAGTCC 413
QY 337 CTTTCCTTCAACACCTTGCATCGGTTTACTTGTGCACTGGACTGCTAGTCCGCTTAT 396
Db 414 CTTTCCTTCAACACCTTGCATCGGTTTAMTTGTGCACTGGACTGCTAGTGTCTGCTTAT 473
QY 397 GCTTATTTGTCGACGATAG 415
Db 474 GCTTATTTGTCGACGTTG 492

RESULT 13

AAC47974
ID AAC47974 standard; DNA; 2660 BP.
AC AAC47974;
XX AC
XX AC
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55799.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000BP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.

PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161320P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.	Query Match 22.6%; Score 270; DB 3; Length 2660;		
PR	05-AUG-1999;	99US-0147192P.	Best Local Similarity 55.6%; Pred. No. 1.1e-71;		
PR	05-AUG-1999;	99US-0147260P.	Matches 544; Conservative 0; Mismatches 425; Indels 9; Gaps 1;		
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.	Qy	125	AAAGGTGGCTTCGTGCTCCATATACGTAATCTTTGTGCCTCTTTTGCAGCCACTAGCTA 184
PR	09-AUG-1999;	99US-0147493P.	Db	189	AATGGTGGTTCCTTGATCCGTTAAACATCTTTTCTTATTTGGTCAAGCTGCTTCGG 248
PR	10-AUG-1999;	99US-0147935P.	Qy	185	CAATTCTGGGTAGATTGTACTATGAAATGGAGAAATAGACATATGTGGTCAACATTC 244
PR	13-AUG-1999;	99US-0148565P.	Db	249	TTCTTCTTGGTAGGTTTACTATGATGAGGTGGAAACAGTAAATGATGCAACTCTTG 308
PR	16-AUG-1999;	99US-0149368P.	Qy	245	TTCAACTCATTTGGCTTCCCTGTACTGGTTCCTGTTCCGCTTCTTTCTCGAATCAGGCAAC 304
PR	17-AUG-1999;	99US-0149175P.	Db	309	TTCAACGGCTGCTTTTCTTATATCTATATCCCGCTTTTGGTACTTCCGCTTCGG--- 365
PR	18-AUG-1999;	99US-0149268P.	Qy	305	CCAAATCAACAGATACAAATTTTCAGTCAGTCCCTTTCCTTCACCACCCTTTGCAATCGGTTT 364
PR	20-AUG-1999;	99US-0149722P.	Db	366	-----CTAGTGTAGAGTCTTCGGAGAGTTCGTTGTTCACTCAAGTACATTTGTTGATCT 419
PR	20-AUG-1999;	99US-0149723P.	Qy	365	ACTGTGCACTGGACTGTAGTGTCCGCTTATGCTTATTTGTTCTGCAATAGGGTGTGCTCT 424
PR	23-AUG-1999;	99US-0149902P.	Db	420	ATGTTTTTGGCTCGGTGTGATCATTTGCTGGAGATAATATGTTATATCTCTGTTGGACTTTGT 479
PR	23-AUG-1999;	99US-0149930P.	Qy	425	ACTTACAGTCTCTACTTTCTCCCTCATCTTGGCCCTCACAGTTGGCCTTCACCTGCCTTTT 484
PR	25-AUG-1999;	99US-0150566P.	Db	480	ACCTCTCTGCATCGAGGTATTCGCTCATTTTGGCGTACTCAGTTAGCTTTTCAACGCGGTGT 539
PR	26-AUG-1999;	99US-0150884P.	Qy	485	TCTCATATTTCCTTAACTCGCAAAAGTTCACTCCTTTGTAGTAGTCAGTTCTTTGTTCTCC 544
PR	27-AUG-1999;	99US-0151065P.	Db	540	TCCTTATTTTCATCAATGCTCAGAGTTCACTGCTTTGATTTCTCAACTCCGTTGTTCTCT 599
PR	27-AUG-1999;	99US-0151066P.	Qy	545	TCACTGTATCCTCTGCTCTTCTTTGTGGTCAACACTGATTCAGAAAACTCAACTAATGAT 604
PR	30-AUG-1999;	99US-0151080P.	Db	600	TGTCGTTTCCCGCTGCTTTTGATAGTCTCTCAATGATGATGCGGATACTCTCTTCTGGTGTCT 659
PR	31-AUG-1999;	99US-0151303P.	Qy	605	CTAGAGTACAGTATGTGATCGGGTTCATATGTATACCATCGGTGCTTCGCTGGGATTGGAC 664
PR	01-SEP-1999;	99US-0151338P.	Db	660	CCAGGTCTAAGTATATTTGTTGGGTTTGTGTGTACACTTGTGCTGTGCTCTCTATTTCTC 719
PR	07-SEP-1999;	99US-0152363P.	Qy	665	TGTTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTTTTCACGAAGCATACATCTCTCAG 724
PR	10-SEP-1999;	99US-0153070P.	Db	720	TGTTGCTATCTCTTATGCACTTCTCGTTCGAGAAGATTCTGAAGAGGGAGACGTTTTCTG 779
PR	13-SEP-1999;	99US-0153758P.	Qy	725	CAGTCACGACCTTGGCCCAITTTACAGTCTCTAGTTGGAGTTGTGTAGTTCTCATAGGAC 784
PR	15-SEP-1999;	99US-0154018P.	Db	780	TGGTTCCTTGAATGCAAACTACACTTCTTTAGTGGCGACTTGTGTTTCGGTTATAGGCG 839
PR	16-SEP-1999;	99US-0154039P.	Qy	785	TTTTTTCAGTGGAGAGTGGGAAACTTTGCAAGTGAGATGAGAAACTACAACTCGGGA 844
PR	20-SEP-1999;	99US-0154779P.	Db	840	TTTTTGTAGCGGGAATGGAGAACGCTGTCATGGGAAATGGAGGTTATCATTAAGGGC 899
PR	22-SEP-1999;	99US-0155139P.	Qy	845	AAGTGTATATGTTTGGACTTTAGCCCTCGGAGCTATTTCTCGCAAGTCTACACTCTTG 904
PR	23-SEP-1999;	99US-0155139P.	Db	900	AAGCTCTTATGTACTGACCTTGTCTGGACAGCAGTTACTTGGCAAGTGTGTTCTCTTG 959
PR	24-SEP-1999;	99US-0155486P.	Qy	905	GTCTTGTGGGATTGATCTTCGAGTCATCTCTCTGTGTCTTCCAATTCATTAACAGCTGCG 964
PR	28-SEP-1999;	99US-0156458P.	Db	960	GAGTCGTGGGTTTGATATTTCTGGTGACGCTGCTCTTCTCAAAACGTCATTAGTACGCTCT 1019
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-015753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	18-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160770P.			
PR	21-OCT-1999;	99US-0160814P.			
PR	21-OCT-1999;	99US-0160815P.			
PR	22-OCT-1999;	99US-0160980P.			

KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
FN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-IB001105.
PF 22-JUN-2001; 2001WO-IB001105.
PR 22-JUN-2001; 2001WO-IB001105.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
XX WPI; 2003-175290/17.
DR
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 6; SEQ ID NO 4071; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
XX Sequence 2175 BP; 363 A; 701 C; 671 G; 439 T; 0 U; 1 Other;
XX
Query Match 17.9%; Score 213.2; DB 8; Length 2175;
Best Local Similarity 52.9%; Pred. No. 2.8e-54;
Matches 492; Conservative 0; Mismatches 423; Indels 15; Gaps 1;
QY 89 CATCATCATCAGTGTGCTCAACCCGAGACTATTAAGGTGGCTTCGTCTCCATAT 148
DB 98 CAGCAGCTCCGGAAACACGACGACCAAGCCATTTCGATGGGCGGATGGTGCAGTGG 157
QY 149 ACCTAATCTTTGTCTCTTTTCCAGCCACTAGCTACAATTCGGGTAGATTGTACTATG 208
DB 158 ACGTCTTCTTCTCATCCCGGTGAGATCGGCGACACTGCTGGCGAGATACTACTACA 217
QY 209 AAAATGGAGAAATAGACATATATGGTAAACATTTCTTCAACTCATTTGGCTTCCCTGTAC 268
DB 218 CCCAAGGCGCGCAGCAAGTGGATATCGGGTTCGTGCGGACGCGCGCTTCCCGATAC 277
QY 269 TGGTTCCTGTCGGCTTCTTTCTCGAATCAGCGACCCAAATCAACAGATACAATTCA 328
DB 278 TGTCTTCACCCCTCTCTTCTTCTCCCTCGAAGTCAACGTCCTCTCTCTCTCTCTCTCT 322
QY 329 GTCACTCCCTTCTCTTCAACCCCTTGCATCGGTTTACTTGTGCACTGGACTGTAGTGT 388
DB 323 GCACCAACACTCCCATGGCCGAGCTCGCGGTGATATACATCGTCTGGGCTCATCATCG 382
QY 389 CGCTTATGCTTATTTGTCTCAGTGGGTTGCTCTACTTACAGTCTCTACTTTCTCCC 448
DB 383 CCGCGGACGACATGATGTACATGCTGGGCTCAAGTACCTCCCGCTTCGACCTACTCGC 442
QY 449 TCATCTTGGCTCACAGTTGGCTTCACTGCTTTTCTCATATTTCTTAACCTCGCAA 508
DB 443 TCATCTGTGCCAGCCAGCTCGCGTTCAATGCTGCTCTCTCATACGTCTCAACTCCCAGA 502
QY 509 AGTTCACTCTTTGATAGTCAGTTCTTTGCTTCTCTCTCACTGTATCTCTCTCTCTTTG 568

DB 503 AGGTCACTCTCTGATATTCAACTCCGTGCTGCTCACCATGTCCTCGCTCGCTCATCG 562
QY 569 TGGTCAACACTGATTTCAGAAAACTCAACTAATGATCTAGAGTACAGTATGTGATCGGGT 628
DB 563 GAGTCAGCAAGGAGTCTCAGGGGGTCAACCGCGTCTCGGAGGAGAGTATCTGCTCGGTT 622
QY 629 TCATATGTACATCGGTGCTTCCGCTGGGATTGGACCTGTTACTATCTCTGATACAAATGC 688
DB 623 TCGTGTGACGCTGGGGGCGTCTGTCACCTACTGCTGATCCTCGCGCTGATGACGCTCA 682
QY 689 TCTTCAGGAAAGTTTTCACGAAGCATACATCTCAGCAGTCAAGGACTTGGCCATTTTACC 748
DB 683 CCTTCGAGACCATCATCAAGAAGCACACCTTCTCAGCCGCTCTCAACATGCAGATCTACA 742
QY 749 AGTCTCTAGTTGGAGTTGTGTAGTTCTCATAGACCTTTTGGCAAGTGGAGAGTGGGAAA 808
DB 743 CGGCGCTCGTGGGACCGCGCGTGGTGGTTCGGGCTGTTTCGGAGCGGAGTGGAGGT 802
QY 809 CTTTGCACAGTGAGATGAGAAACTACAAACTCGGGAAGTGTATATGTTTGTGACTTTAG 868
DB 803 CGCTGAGGGGAGATGAACGCTTACGCTGGGGGAGTTCTCTACTCTGATGACGCTGC 862
QY 869 CCTCGGCGAGCTATTTCTCGCAAGTCTTACACTCTCTGCTCTTGTGGGATTTGATCTTCGAGT 928
DB 863 TGTGGGCGCGGTGCTGTCGAGGTGGCCAAACATCGGGGTGCTCGGCTCATCTTCGAGG 922
QY 929 CATCTCTGTGTTCTCCAAATTCATACAGCTGTGGGATTCCTTATAGTTCAGTTCGCGG 988
DB 923 TGTGGGCGCTTTCTCCAAAGTGTATCAGCACGCTGTCACTGCGCGTCACTCCCGTCTCTTCG 982
QY 989 CAGTGTAGTTTTCATGATAGTGAACG 1018
DB 983 CGGTGGTGGTGTTCACGACAGGATGAACG 1012

Search completed: November 1, 2004, 07:20:01
Job time : 632.985 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:45:29 ; Search time 4079.18 Seconds
(without alignments)
10228.408 Million cell updates/sec

Title: US-09-913-767-3
Perfect score: 1145
Sequence: 1 ctgttctatgtgttgatggt.....caagtttaggcaaatgacca 1145

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1051.4	91.8	1323	3	CNS0A36P
2	316.8	27.7	1431	3	Arabidopsis
3	315.6	27.6	1340	3	Arabidopsis
4	312	27.2	1375	3	Arabidopsis
5	307.8	26.9	1312	3	Arabidopsis
6	287.2	25.1	1359	3	Arabidopsis
7	252.4	22.0	1374	3	Arabidopsis
8	246.6	21.5	782	7	Arabidopsis
9	239	20.9	1519	3	Arabidopsis
10	231.4	20.2	922	8	Arabidopsis
11	229	20.0	770	9	Arabidopsis
12	225	19.7	768	1	Arabidopsis
13	220.8	19.3	739	7	Arabidopsis
14	219.4	19.2	837	6	Arabidopsis
15	218.2	19.1	509	1	Arabidopsis
16	217.2	19.0	666	5	Arabidopsis
17	217	19.0	1299	3	Arabidopsis
18	215.6	18.8	1289	3	Arabidopsis
19	213.8	18.7	633	1	Arabidopsis
20	204.6	17.9	905	5	Arabidopsis
21	203.8	17.8	627	5	Arabidopsis
22	200.2	17.5	630	4	Arabidopsis
23	200.2	17.5	762	5	Arabidopsis
24	195	17.0	735	4	Arabidopsis

25	194	16.9	806	4	BI310188
26	190.2	16.6	762	4	BI413141
27	183.6	16.0	568	4	BM176948
28	183	16.0	701	4	BI204933
29	182.6	15.9	699	8	B77316
30	182.6	15.9	800	6	CF069001
31	180.8	15.8	541	4	BI973110
32	179.8	15.7	676	6	CA920055
33	179.2	15.7	742	4	BI205696
34	179.2	15.7	742	4	BI210251
35	178.2	15.6	882	7	CNI25340
36	177.6	15.5	591	4	BI205616
37	174.8	15.3	593	4	BI432746
38	174.6	15.2	589	7	CK096760
39	174.6	15.2	765	7	CNI46235
40	174.4	15.2	753	4	BM409055
41	173.6	15.2	691	8	BZ087060
42	173	15.1	757	5	BQ515368
43	172.3	15.1	761	7	CNI25253
44	172.2	15.0	562	7	CK759270
45	171.8	15.0	480	4	BG652290

ALIGNMENTS

RESULT 1	CNS0A36P	1323 bp	mRNA	linear	HTC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTBF642F05 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).				
DEFINITION	Arabidopsis thaliana (thale cress)				
ACCESSION	BX826811				
VERSION	BX826811.1	GI:42460388			
KEYWORDS	HTC; GSLT cDNA.				
SOURCE	Arabidopsis thaliana				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.				
TITLE	Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1323)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	The sequences are based on single pass reads. Life technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Pull_length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.				
FEATURES	Location/Qualifiers				
source	1..1323				
	/organism="Arabidopsis thaliana"				
	/mol_type="mRNA"				
	/strain="Col-0"				
	/db_xref="taxon:3702"				

BI310188	EST531193
BM413141	EST587468
BM176948	aej74h01
BI204933	EST522973
B77316	T3213TF TAM
CF069001	EST669722
BI973110	sa440a05
CA920055	EST637773
BI205696	EST523736
BI210251	EST528291
CNI25340	RHOH1.10
BI205616	EST523656
BI432746	EST535507
CK096760	UB25CPF10
CNI46235	WOUND1.39
BM409055	EST583382
BZ087060	lkh1od1.4
BQ515368	EST622783
CNI25253	RHOH1.10
CK759270	pam01-13m
BG652290	sa665a10

Qy	28	GAAGGAAATTTTCAACAGAGAGAGAAAGTCAAGTACTCTTGGAGGTTAAGAGTGTCT	87
Db	197	GAATCTGAATCAITTTTCAGTACCTCAACGAGAACTGTAAGAGGTGCTCCGTGTCTCC	256
Qy	88	CTCTATGTCACTCTCTCTAGCTGGAGAGACAATAGCACCTCTCTAGGTAGACTTTAC	147
Db	257	ATATACGCAATCTTTGTICATCTTCTGCCAACCATCTGCTCAGTCTCTGGGTAGACTGTAC	316
Qy	148	TACGAAAAAGCGGTAAAGCACATGGCTCGAAACCTTTGGTTTCAGCTTGTAGGGTTTCCT	207
Db	317	TATGAAATGAGGAGAAAGCACATATGTGGTAAACACTTCTTCAACTCATTTGGCTCCCT	376
Qy	208	TTAACCCCTCTCTGCTATTAATCTTAAAGCTGAGCGTCAAGACTAAACCAATTACC	267
Db	377	GTACTGATCTCTTCGCTCTCTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAT	436
Qy	268	AAAAAACTACTCTCTCTCTTGACACTATCTTTAGTGATATATGGACTTGGCTTGCTT	327
Db	437	TTCTAGTCAGTCCCTTCTTCCACACCCTTGCAATCGGTTTACTTGTGCACGTGGACTGCTA	496
Qy	328	GTTGCTGGACATGTATTTTGTACTCATTTGGGCTACTTTTACCTTCTCTGTCTCAACTTTC	387
Db	497	GTGCTGCTATGTCTTATTTGCTGCGAGTGGGTGCTTTACTTACCACTCTCTACTTTC	556
Qy	388	TCTTTGATCTCTGGCTGCGAATTTGGCTTTTAAAGCCGTCCTCTTACTTCTTAAACTCA	447
Db	557	TCCCTCATCTTGGGCTCACAGTTGGCCTTCTACTGCTTTTTCTCATATTTCTTAACTCG	616
Qy	448	CAAAAATCACACCATTTATATCTCAATTCACCTTGTCTTCTTAAACATATCTTACACTT	507
Db	617	CAAAAGTTCACTCTCTTTGATAGTCAATCTTTGTTTCTCTTACGGTTTCTCTGCCCCC	676
Qy	508	CTTGTTATCCAACTGAACCAAGAAATCTCCCTCTTCTACTTCAAAAGTCGACGCAAGTCC	567
Db	677	CTCGTGTCACTGATTCAGAA-----AACACAAACAATGATCTAGAGTA	724
Qy	568	AAGTATGTGATTTGATACATCTGCGCGGTGCGTAGCTCAGCTGGTTATTTCTCGTGCTT	627
Db	725	CAATATGTGATTTGGGTTTCACTGTACCAATGGTGTCTCCGCTGGGATTTGGATTTGTA	784
Qy	628	TCTTTAACAGATTAGCGGTTTCGAAAGATTTCTAAAGAAATACACATTTCAAGGCTATT	687
Db	785	TCTCTGATACAACTGCTCTTCAGGAAAGTTTTCAGAAAGCATATCTCTCAGCAGTCTG	844
Qy	688	GACATGGCCACATATCCGTCTATGTTAGTACTTGTGTAGTTGTGGTAGGACTTTTTTGG	747
Db	845	GACTTGGGCAATTAACAGTCTCTAGTTGCTACTTGTGTGGTACTCATAGGACTGTTTGA	904
Qy	748	AGTGGTGGTGGAAAAAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGGAAAAAGCTCA	807
Db	905	AGTGGAGGTGGAGAAATCTCCCAAGTGAGATGAGAAACTACATACTGGGGAAGTGCA	964
Qy	808	TACATTTTGATPAAACATCGGTTTCAACGATATCATGGCAGCTTGTGTGATTGGAAGTGT	867
Db	965	TATATCTTGACTTTTGGGCTCAGCAGCTATTTTCTGGCAAGTATACACTGTTGTTGTGTG	1024
Qy	868	GGTTTGATATTCGAAGTTTCACTCGCTTTTTTCCAAATGCTAAAGCATCTTTTGTATTACA	927
Db	1025	GGATTAATCTTCAGTCTTCTCTGTGTTTCTCCAAATTCATPACTGCTGTGGGACTGCTT	1084
Qy	928	GTTGTGCTGTTCTTGTCTGTTGTCTTCTTCGCTGATGAGATGAGTGAATCAAGTTGCTT	987
Db	1085	ATAGTTCCAGTTGTAGCAGTGATAGTTTTCATGATGAAGTGGATGCATCAAGATTTTC	1144
Qy	988	GCAATGTTTGTGGCATCTGGGAAATTTGTTTCTTATGTTATACAGATATATGTCAATGAT	1047
Db	1145	TCCATCATTTTAGCTATCTGGGGCTTCCTTTTCAITTTGTCTATCAGAACTACCTCGGAA	1204
Qy	1048	AGAAAG 1053	
Db	1205	AAGAAG 1210	

RESULT 3	CNSO29J	1340 bp	mrna	linear	HTC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSL7S171ZA05 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).				
ACCESSION	BX827536				
VERSION	1	GI:42459416			
KEYWORDS	HTC; GSUT cDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
AUTHORS	Castelli, V., Aury J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.				
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1340)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externes/sequences/Banque_Projet_EF/Full-length				
FEATURES	source				
gene	1..1340				
ORIGIN	Query Match 27.6%; Score 315.6; DB 3; Length 1340; Best Local Similarity 57.4%; Pred. No. 3e-73; Matches 595; Conservative 0; Mismatches 429; Indels 12; Gaps 1;				
Qy	18	GGTAGAACCTGGAAGGGAATTTTCAACAGAGAGAGAAGTCAACAGTACTCTTTGGAGGTT	77		
Db	103	GGTGGTAACCTGAATCATCATCATCAGCTGTGCGCTCAAAACCGAGAACTATAAAAGGTGGCT	162		
Qy	78	AAGAGTGTCTCTTANGTCACTCTCTCTTAGCTGGAGAGACAATAGCCACTCTCTTAGG	137		
Db	163	TCGTGTCTCATATACGTAACTCTTTGTGCTCTTTTTCGCCAGCCACTAGTACAATCTTGGG	222		
Qy	138	TAGACTTTTACTGAAAAAGCGGTAAAAAGCATGGCTCGAAACCTTGGTTCAGCTTGT	197		
Db	223	TAGATTGTACTAGAAATGGAGAAATAGCACATATGGTAACACTTCTTCACTCAT	282		
Qy	198	AGGGTTTCCTTTAACCCCTTCCTTGCCTATTATTACTTTAAAGCCGTGAGCCGTCGAAGACTAA	257		
Db	283	TGCGTTCCTCTGTAAGTGTCTGTTCGCTCTCTTTCTCGAATCAGGCAACCCAAATCAAC	342		

Qy	258	AACCATTAACAAAAA	CTACTCTTCTTGACA	CTATCTTTAGTGATAT	GGACT	317
Db	343	AGATACAAATGTCAGT	CAGTCCCTCTCTTCA	CCACCCTTGCAT	FCGGTTTACTTTGTGCAC	402
Qy	318	TGGCTTGCTGTGTGTCG	GACATGATATTTTGT	ACTCATTTGGGCTACT	TTTACCTTCCTGT	377
Db	403	TGGACTGCTAGTGTCCG	CTTATGCTTATTTGT	CTGCAGTAGGGTGTCT	TACTATTACCAGT	462
Qy	378	CTCAACTTCTCTTTGAT	CTCTCGGTCGCAAT	TGGCTTTTAAAGCCGCT	CTTCTCTTACTT	437
Db	463	CTCTACTTTCTCCCTCA	TCTTGCCCTCACAGT	TGGCCCTTCACTGCCCT	TTTCTCATATTT	522
Qy	438	CCTAAACTCACAAAAAT	CACACATTTATTACT	CAATTCATTTGTTCTCT	TTTAAACCATATC	497
Db	523	CCTTAACTCGCAAAAGT	TCACTCCTTTGATAGT	CAGTTCTTTGCTTCT	CTCTCACTGTATC	582
Qy	498	TTCTACACTTCTTGTTAT	CCAAACATGAACAGAA	TCTCCCTCTTCTACTT	CTTCAAGTCCGC	557
Db	583	CTCTGCTCTTCTTGTG	TGTCACACTGATT	CAGAA-----AACT	CAACTAATGT	630
Qy	558	AGCCAAGTCCAAGTATG	TGATTTGGATACATCT	TGCGCGTTCGTTAGCT	CAGCTGGTTATTC	617
Db	631	ATCTAGAGTACAGTATG	TGATCGGGTTTCATAT	GTATACCATCGTGTCT	CCGCTGGGATTTG	690
Qy	618	TCTGGTGTCTTCTTTAA	CAGATTACGGGTTT	CGAAAGATTCTAAAGAA	ATACACATTCAA	677
Db	691	ACTGTTACTATCTCTGA	TACAAATGCTCTTTC	AGGAAAGTTTTCAGAA	GACATACATCCTC	750
Qy	678	GGCTATTTTAGACATGG	CCACATATCCGTTCTA	TGGTAGTACTTTGTGTAGT	TTGTGTGGTAGG	737
Db	751	AGCAGTACAGGACTTGG	CCATTTTACCAGTCTCT	AGTTGCGAGTTGTGTAGT	TTCTCATAGG	810
Qy	738	ACTTTTGGAAAGTGGTG	GGTGGAAAAAGCTG	AGTACAGAAATGGAAGAGT	TTTCAACTAGG	797
Db	811	ACTTTTTCGAAGTGGAG	ATGGGAACTTTGCCA	AGTGAGATGAGAACTAC	AAACTCGG	870
Qy	798	GAAGAAGTCATACATTT	TGTATAAACATCGGTT	CAACGATATCATTGGCA	AGCTTGTTTGAT	857
Db	871	GAAGAAGTGCATATGTT	TGTACTTTAGCTTCGGC	TAGCTTATTTCTTGGCA	AGTCTACACTCT	930
Qy	858	TGGAAGTGTGTGTTGAT	TATCGAAGTTTTCATCG	CTTTTTCCTCAATGTCA	TACAGCACTCT	917
Db	931	TGTCCTGTGGGATGATCT	TCGAGTCATCTCTGTGT	TTCTCCAAATTCCTATA	TACAGCTGT	990
Qy	918	TTGCTTTACCAGTGTGCT	GTCTTCTTCTGTTGTCTT	CTTCGTCATCAGATGAGT	TGGAAT	977
Db	991	GGGATTGCCCTATAGT	TCAGTTGCGGCAGTGAT	AGTTTTCATGATAGAA	TGGACGCATC	1050
Qy	978	CAAGTTGGTTGCAATGTT	TTTGGCCATCTCGGGAT	TTGTTTCTTATCGTTAT	CAGCATTA	1037
Db	1051	AAAAATCTTCTCCATAT	TTTATAGCTATCTCGGCT	TCTCTCATTCGTTAT	CAGCACTA	1110
Qy	1038	TGTC	CAATGATAGAAAG	1053		
Db	1111	CCTC	GACGAAAGAAAG	1126		
RESULT 4						
CNS0A2S8						
LOCUS						
DEFINITION	CNS0A2S8 1375 bp mRNA linear HTC 06-FEB-2004					
Arabidopsis thaliana Full-length cDNA Complete sequence from clone						
GSLTUS202A10 of Adult vegetative tissue of strain col-0 of						
Arabidopsis thaliana (thale cress).						
ACCESSION	BX827187					
VERSION	BX827187.1 GI:42459879					
KEYWORDS	HTC; GSLT cDNA					
SOURCE	Arabidopsis thaliana (thale cress)					
ORGANISM	Arabidopsis thaliana					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;						
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.						
REFERENCE	1 (bases 1 to 1375)					

AUTHORS		Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.	
TITLE		Whole Genome Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation	
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1375)		
AUTHORS		Genoscope.	
TITLE		Direct Submission	
JOURNAL	Submitted (18-NOV-2003)	Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)	
COMMENT		- Web : www.genoscope.cns.fr) The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=arabidopsis.	
FEATURES		Location/Qualifiers	
source		1. .1375 /organism="Arabidopsis thaliana" /mol_type="mRNA" /strain="Col-0" /db_xref="taxon:3702" /clone="GSLT202A10" /tissue_type="Adult vegetative tissue" /plasmid="pCMVSPORT_6" 1. .1375 /gene="At4g18205"	
gene			
ORIGIN		Query Match 27.2%; Score 312; DB 3; Length 1375; Best Local Similarity 57.4%; Pred. No. 2.8e-72; Matches 589; Conservative 0; Mismatches 425; Indels 12; Gaps 1;	
Qy	28	GAAGGGAAATTTCAACAGAGAGAGAGTACCAAGTACTCTTGGAGTTTAAAGAGTCTCT	87
Ds	111	GAACAGTGAATCATTTTCAGTACCTCAAACGAGAACTGTAAAGGTGGCTCCGGTCTCC	170
Qy	88	CTCTATGTCACTCTCTCTTAGCTGGAGAGACAATAGCCACTCTCTTAGGTAGACTTTAC	147
Ds	171	ATATACCATCTTTGTCATCTTCTGCCAACCACTTGTCTACAGTTCTGGGTAGACTGTAC	230
Qy	148	TAGCAAAAGCGGTAAAGACATAGGCTCGAAACCTTGGTTCAGCTTGTAGGTTTCCT	207
Ds	231	TATGAGAAATGGAGGAAAGACATATGTGTAACTTCTTCAACTCATTTGGCTTCCCT	290
Qy	208	TTAACCTTCTTGTCTATTATTACTTAAAGCCTGAGCGTCCAGAGATTAACCACTACC	267
Ds	291	GTACTGATTTCTTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAAT	350
Qy	268	AAAAAACTACTTCTTCTTGTGACACTATCTTTAGTGATATTTGGACTTGGCTTGCCTT	327
Ds	351	GTCACTGAGTCCCTTCTTCCACCACCTTGCATCGGTTACTTGTGCACCTGGACTGCTA	410
Qy	328	GTGCTCGCAATGTATTTGTACTCATTTGGGCTACTTTACCTTCTCTGTCTCAACTTTC	387
Ds	411	GTGCTCTGTTATGCTTATTTGTCTGCAAGTTGGGTGCTTTACTTACCAGTCTCTACTTTC	470
Qy	388	TCCTTGTATCTCGTTCGCAATTTGGCTTTTAAACGCGCTTCTTCTTACTTCTTAAACTCA	447
Ds	471	TCCCTCATCTTGGCTCACAGTTGGCTTCTCACTGCCCTTTTCTCATATTTTCTTAATTCG	530
Qy	448	CAAAAAATCACCAATTTATCAATTTCACTTGTCTTCTTAACCATATCTTCTTACACTT	507

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium. 1 (bases 1 to 782)

REFERENCE
AUTHORS

Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and Wing, R. A.

TITLE
JOURNAL

Global assembly of Cotton ESTs

COMMENT

Unpublished (2004)

Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: <http://genome.arizona.edu>

Plate: 0035 row: G column: 15.

FEATURES

source

1..782

/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_EB0035G15"

/tissue_type="floral"

/dev_stage="3 to +3 DPA"

/lab_host="DH10B"

/clone_lib="GR_Eb"

/note="Vector: pCMV.SPORT-6.1; Site1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 21.5%; Score 246.6; DB 7; Length 782;
Best Local Similarity 64.1%; Pred. No. 8.4e-55;
Matches 394; Conservative 0; Mismatches 209; Indels 12; Gaps 1;

Qy 444 CTCACAAAATCACACCATTTACTCAATTCACCTGTTCTCTTAACCATATCTCTAC 503
Db |||||
Qy 782 CTCACAAAGTTCACCCCTTTTCATAATAACTCTTTGGTCTCTCCATCAATCTCTCCAC 723
Db |||||
Qy 504 ACTTCTTGTTATCAACATGAACAGAAATCTCCCTCTTCTACTTCAAAAGTCCGACGCCAA 563
Db |||||
Qy 722 CTCTCTGATTCAAAACGAT-----TCCTCAGTTCACACAGTTCACAG 675
Qy 564 GTCCAAAGTATGTGATGTGATACATCTCGCGGTGGTAGCTCAGCTGTTATTTCTGTGT 623
Db |||||
Qy 674 AGCCAAATATGTATCGGATTTGTATGACCGGTGGCTGCTTCAGCAGGGTACGGATTGAT 615
Qy 624 GCTTCTTTTAACAGATTACGGTTCGAAAGAGTTCTAAGAAATCTAAGAAATACATTCAGGCTAT 683
Db |||||
Qy 614 GTTGCTCTTAACACAACTTTGTTTCCAAAAGATTCCTCAAGAAACAATCATTTAAGGCAGT 555
Qy 684 TTTAGACATGGCCACATATCCGTCTATGGTAGTACTTGTGTAGTTGGTAGGACTTTT 743
Db |||||
Qy 554 GTTGACATGATATCTACCATGTCATCGCTGCACTTTTGATGATCATGCTGGGTCTTTT 495
Qy 744 TGGAAAGTGGTGGGAAAGAGCTGAGTACAGAAATGGAAGAGTTTCACTAGGGAAG 803
Db |||||
Qy 494 GCGCAGTGGGAAATGGAAGAGTCTGGGCGAAGAAATGGAAGGTTTAACTGGGTGAAGTC 435
Qy 804 CTCATACATTTGATAAATCATCGTTCAACGATATCATGGCAAGCTTTGTTGATGGAG 863
Db |||||
Qy 434 GGCATACGTTAACGTTTTGGTTGGATTGCTGATGGGTGGCAAGTTTTCTCGATTGGTGC 375
Qy 864 TGTGTTGTTGATATCGAAGTTTCATCGCTTTTTTCCAAATGTCATAAGCACTTTTGTGTT 923
Db |||||
Qy 374 GGTGGTTTGTATCTTCGAAGCGTTCATCGCTTTTCTCAATATGTATCAGCAGTTGGGACT 315
Qy 924 ACCAGTTGTGCTGTTTCTGTTGTTCTTCTTCGTTGATGATGATGATGATGATGATGATGAT 983
Db |||||
Qy 314 ACCCAATGTTCCAGTGTGTCGATGTTGTCGATGTTTTCACGATCCAAATGACTGGAATTAAGGT 255

Qy 984 GGTTCGAATGTTTGGCCATCTGGGATTTGTTCTTATGTTATCAGCATTTATGCAA 1043
Db |||||
Qy 254 CATTTCCATGTTGTTGGCCATATGGGATTTTGCTCTATGTTTACCAGCACTATCTGA 195
Db |||||
Qy 1044 TGATAGAAAGCCAGA 1058
Db |||||
Qy 194 TGATCGAAACTCTGA 180

RESULT 9

CNS0A3B5

LOCUS

DEFINITION

Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFB632D12 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION

Version BX826795.1 GI:42460970

KEYWORDS

HTC; GSLT cDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS

1 (bases 1 to 1519)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 1519)

AUTHORS

TITLE

JOURNAL

COMMENT

The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length

FEATURES

source

Location/Qualifiers

1..1519

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="Col-0"

/db_xref="taxon:3702"

/clone="GSLTFB632D12"

/tissue_type="Flowers and buds"

/plasmid="pCMVSPORT_6"

1..1519

/gene="At4g18210"

ORIGIN

Query Match 20.9%; Score 239; DB 3; Length 1519;
Best Local Similarity 56.1%; Pred. No. 1.1e-52;
Matches 600; Conservative 0; Mismatches 445; Indels 24; Gaps 7;

Qy 57 TCACAAGTACTCTGGAGGTTAAGAGTCTCTCTATGTCACCTC-TCTCTTAGCTGAG 115
Db |||||
Qy 194 TAACACATACAAACGGTGGCTCAGGGTACATTATATACATTTCTTTGTCAATTCAGGCC 253
Db |||||
Qy 116 AGACAATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAAAGCGGTAAGCAATGCGC 175
Db |||||

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 06:20:04 ; Search time 559.891 Seconds
(without alignments)
9900.026 Million cell updates/sec

Title: US-09-913-767-6

Perfect score: 1081

Sequence: 1 caaatccacagttcaagat.....ataagtgatatttgaag 1081

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	406.2	37.6	3387	9	US-09-938-842A-2212
2	406.2	37.6	3387	11	US-09-938-842A-2212
3	361.8	33.5	1294	16	US-10-424-599-2524
4	307	28.4	1513	17	US-10-437-963-82061
5	303.8	28.1	2214	16	US-10-424-599-102519
6	262	24.2	2381	17	US-10-437-963-43860
7	234.2	21.7	83698	17	US-10-416-898-9
8	223.4	20.7	997	16	US-10-424-599-75161
9	179.2	16.6	728	17	US-10-767-701-9000
10	149	13.8	522	16	US-10-424-599-114931
11	141.6	13.1	1386	16	US-10-425-114-12485
12	138.6	12.8	592	9	US-09-770-152-44
					Sequence 2212, Ap
					Sequence 2212, Ap
					Sequence 2524, Ap
					Sequence 82061, A
					Sequence 102519, A
					Sequence 43860, A
					Sequence 9, Appli
					Sequence 75161, A
					Sequence 9000, Ap
					Sequence 114931, A
					Sequence 12485, A
					Sequence 44, Appli

13	138.4	12.8	1779	16	US-10-425-114-28473
14	131.8	12.2	1233	16	US-10-424-599-30274
15	131	12.1	1146	16	US-10-425-114-10516
16	129.2	12.0	2605	17	US-10-437-963-83900
17	127	11.7	1432	17	US-10-437-963-1897
18	126.4	11.7	332	11	US-08-732-627A-1196
19	125.8	11.6	671	17	US-10-767-701-5986
20	123	11.4	1373	16	US-10-424-599-31364
21	117.6	10.9	792	16	US-10-424-599-9213
22	115.2	10.7	1333	16	US-10-425-114-10574
23	109.6	10.1	1355	16	US-10-424-599-11363
24	108.2	10.0	327	16	US-10-424-599-121152
25	102.6	9.5	650	17	US-10-767-701-4045
26	97.4	9.0	747	17	US-10-767-701-6806
27	82.2	7.6	994	16	US-10-424-599-73501
28	73.8	6.8	308	16	US-10-424-599-121805
29	72.2	6.7	774	9	US-09-770-445-895
30	71.4	6.6	1173	17	US-10-437-963-57539
31	66.2	6.1	386	9	US-09-770-791-85
32	63.2	5.8	716	16	US-10-424-599-61310
33	60.4	5.6	1417	16	US-10-424-599-97043
34	59.6	5.5	2928	17	US-10-437-963-33800
35	57	5.3	789	17	US-10-437-963-72452
36	56.4	5.2	1177	17	US-10-437-963-91425
37	55.4	5.1	382	17	US-10-021-323-592
38	55.4	5.1	383	17	US-10-021-323-2053
39	42.6	3.9	59475	17	US-10-322-696-166
40	42.2	3.9	201	17	US-10-741-601-16181
41	42.2	3.9	81210	17	US-10-741-601-5664
42	42	3.9	357652	17	US-10-322-696-34
43	41.8	3.9	1296	16	US-10-425-114-13687
44	41.4	3.8	684187	17	US-10-367-094-71
45	40.6	3.8	72604	15	US-10-162-497-7

ALIGNMENTS

RESULT 1
US-09-938-842A-2212
; Sequence 2212, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2212
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2212

Query Match 37.6%; Score 406.2; DB 9; Length 3387;
Best Local Similarity 64.7%; Pred. No. 2.2e-104;
Matches 622; Conservative 0; Mismatches 333; Indels 6; Gaps 1;

QY 88 GGAGTATCTATGCAACGGAGGAAACAGTAATAGCTAGCAACGGTAGTTCACCTTGT 147
DB 236 GTAGTTTACTATGAAATGGTGGGAAGAGATCATGATGGGGAACACTTGTTCACACTATC 285

Db 946 CTGATCTTTGAGTCATCTTCTGTGTTCTCCTCAATTCATTAACCTCTCTGGGATTCGCTATA 1005

Qy 862 GTTCTCTATCTTGCTGTGTAATCATTTTCCATGACAAATGAATGGGTAAAGGTGATTTCT 921

Db 1006 GTTCCAGTTGTAGCAGTGATTTGTTTCCATGATAAAATGAACCGTCAAAGATCTTCTCC 1065

Qy 922 ATGATCCTAGCTATTTGGGGTTTCACTTCTCTATGTTCTACCAACAATATCTTGATGACAAA 981

Db 1066 ATCATTTTAGCTATCTGGGGATTCATTTCTATTTGTTCTATCAGCACTACCTCGACGAAAAG 1125

Qy 982 AACTTGAAGAAAAATCATGAAATCACAAACAAGAAATCCCTGACCCACAGAGCAGAA 1041

Db 1126 AAGTTGAAGACTAGCCACACAGTCTCTGTAGGAGATCCTCATCTACTACTCTGCTAGGAA 1185

Qy 1042 G 1042

Db 1186 G 1186

RESULT 3

US-10-424-599-2524

; Sequence 2524, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; NUMBER OF SEQ ID NOS: 2003-04-28

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 2524

; LENGTH: 1294

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_102285C.1

US-10-424-599-2524

Query Match 33.5%; Score 361.8; DB 16; Length 1294;

Best Local Similarity 63.7%; Pred. No. 5.6e-92;

Matches 596; Conservative 0; Mismatches 312; Indels 27; Gaps 2;

Qy 87 AGGAGTATCTATGACAAACGGAGGAACAGTAGTAATGGCTAGCAACGGTAGTTCAACTTCT 146

Db 199 AGATTGTACTAGAAAGGGTGGAAAAAGCAAGTGGATGGGAACACTTGTTCACACTTGC 258

Qy 147 TGGCTTTTCCCTGCTACTTCCATATTATATCTTTGTCAATTTAAACACATGCAACCACTGA 206

Db 259 TGGTTTCCCTATTTCAGCTTCTTTCACITTTATTTCCAGCATCCAAAATCTCACCAAAA 318

Qy 207 TAG-----AGATGGAAGAAAGACCTCACCTAGGAACCGGTGATTTGGTTTAGCTAGTGT 260

Db 319 TAGTAGCATTCATCCAAAACAATCATCTGCTTCAATACTAGCAATTTATCTATGCTCAAT 378

Qy 261 TGGACTTCTTGTAGGAGCAGATTGCTATCTGTACTCCATTGGACTTCTTTACTTACCCGT 320

Db 379 TGGCCCTACTTTTGGCAATTAGATTGTTATTTATATTTTATATTCAGTTGGACTATGGGACTTCCCTGT 438

Qy 321 TTCTACTATTCCCTGATCTGTGCATCTCAGTTAGTCTTCAATGCTTCTTCTCTTATTT 380

Db 439 CTCTACTATTCACTCATTTGCTCATCCCAATTTGGCTTTCAATGCTTCTTCTTCTTCTTCT 498

Qy 381 TCTTAATCACAAAACTTACCCCTATCATTTTAAATTTCTTTTCTTCTTCTTCTTAACTATATC 440

Db 499 CCTCAACTCACTCAAGTTCACACCTTACATAATCAACTCTCTAGTCTCTTCTCACCATTTC 558

Qy 441 TTCCACCCCTACTTGCATTTCAATAATGAG-----GAGACAGACTC 479

Db 559 TTCAACCCCTCTTGTGTTTCAAAATGAGTCATCATCAGATGATGATGATTCAGATTTC 618

Qy 480 CACAAAAGTTACAAAAGGAGAGTAGTATCTCAAAGTTTTCATATGCACCGTTGCTCGCTCTGC 539

Db 619 CACAAAATCTCCAAAGAAAAGTAGTGTGATTGGATTTCATATGCACAGTTGGTGCACTCTGC 678

Qy 540 TGGTTATGGTCTAGTCTTATCCCTACACAGCTAGCTTCTTAAAGAGTCTTAAAGAGCA 599

Db 679 AGGGTATGGACTATGGCTTTCCCTCACACAGCTTGTGTTCAAAGAGGTCATAAAAAGAGA 738

Qy 600 AAATTTCTCAGAAAGTTATGGATATGATTAATCTACGTGAGTCTAGTGGCCAGTTGTGTTAG 659

Db 739 AACATTCAAAGTGGTCTTGGATATGATTAATATACACTTCCCTTGTGGCTACCCCTTGTCTAC 798

Qy 660 CGTGTGGGGCTTTTGTGTAGCAGTGAAGTGGAAAACTTTTGAGCAGTGAATAAGTAACTA 719

Db 799 CTTAGTGGGACTTTTGTCTAGTGGAGAGTGGAGTGGTTTGAAGAATGAAATCAAGGAGTA 858

Qy 720 CAAACATGGGAAGGTATCTCATATTAAGAACCTAGTGTGGACAGCTGTTACCTGGCAGTT 779

Db 859 TGAGTTGGGGAAGGCTTCATATTTCTGAAACCTCATTTTCACAGGCATACATTTTGGCAAGT 918

Qy 780 ATTTCTCCATCGGTGGCAGGACTGATCTTCGAGCTCTCTCTCTATTTCTCAAAATGCAAT 839

Db 919 CTTACCAATTTGTTGTTGGGCGATATTAGTGAGGTTTCTTCCCTCTCTCTAAATGCCAT 978

Qy 840 AAGCGTTTGGGACTCCAGTGGTTCTTATCTTGGCTGTGTAATCATTTTCCATGACAAAAAT 899

Db 979 AAGCGTTTGGGAGTGCCTATTGTTTCCAATGCTGCGAGTGTGTTCTTTTCATGACAAAAAT 1038

Qy 900 GAATGGGTTAAAGGTGATTTCTATGATCCTAGCTATTTGGGGTTTCACTTCTCTATGCTA 959

Db 1039 GGAATGGCATTAAAGGTTATCTCTATGGTGTGTAGCTATTTGGGGCAATTGTATCATATGTGA 1098

Qy 960 CCAACAATATCTTGATGACAAAAACTTGAAGAAAA 994

Db 1099 TCACAGTACTTGGATGATACAAAGTCCGAAAAACA 1133

RESULT 4

US-10-437-963-82061/c

; Sequence 82061, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 82061

; LENGTH: 1513

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_81525C.1

US-10-437-963-82061

Query Match 28.4%; Score 307; DB 17; Length 1513;

Best Local Similarity 59.4%; Pred. No. 2.5e-76;

Matches 545; Conservative 0; Mismatches 360; Indels 12; Gaps 1;

Qy 92 TATACTATGACACCGGAGGAAACAGTAATCGCTAGCAACGGTAGTTCAACTTGTGGCT 151

Db 1185 TCTACTCAATCAAGGCGGCAATAGCAAGTGGATGCCATTCGTCCAAACTGCTGGCT 1126

```
Qy 152 TTCC-----TGTCCTACTTCCATATATATCTCTGCTATTAATAAACACATGCAA 199
Db 1125 TTCCGATTTTGTTCATTCGCCCTATTTCTTTTCCATTCAAAGACATCTTCTACAAACAG 1066
Qy 200 CAACGTATAGAGATGGAAGAAACCTCACCTAGGAACCGTGATATGGTTTACGTAGTGC 259
Db 1065 TCACCTAGTAGCTGCGCCCTACAATTTCTATCCCCAAAATTAATCTGATATATGTGTCC 1006
Qy 260 TTGGACTTCTTGTAGGAGCAGATTCATCTGTAATCTCAATTTGGACTTCTTTACTTTACCCG 319
Db 1005 TGGGCTCATCATTTGCTGACACGACTTGATGTATTTCTTATGGCTACTATATCTTCGG 946
Qy 320 TTCTACCTATTCCTGATCTGTCATCTCAGTTAGCCTTCAATGCTTTCTTCTCTTAT 379
Db 945 TCTCAACATATTCGCTCATCTGTGTCATGCTAGCTTGCCTTCAATGCTGTCTTCTCATAT 886
Qy 380 TTCTTAACCTCAAAAACCTTACCCCTCATCTATTTTAAATTTCTTTTCCCTCTTAATATAT 439
Db 885 TCCCTCAATGCTCAAAAATTCACCCCTCTGATTTTCAATTCGCTAGTCTCTTACGTTTT 826
Qy 440 CTTCCACCTACTTGCATTCATTAATAGGAGACAGACTCCACAAAAGTTTCAAAAAGAG 499
Db 825 CTGCTTCACTCTCTGGAGTTGATGAATTTCTCAGGGAATCTACTAGTATATACATGGGA 766
Qy 500 AGTATGTCAAAAGTTTCATATGCACCGTTGCTGGCTGCTGGTTATGGTCTAGTCTTAT 559
Db 765 AGTACATTTTGGGTTTCTGTGTGACACTAGGGGATCAGCTACATCTCGCTCATCTCT 706
Qy 560 CCCTACACAGCTAGCCTTTCTAAAGTCTTAAAGAACCAAAATTTCTCAAGAGTTATGG 619
Db 705 CCCTGATGCAAGTACATTTGAGAGGTTATTAAGAGGGAGACCTTCTCAGTTGTGTGA 646
Qy 620 ATATGATATCTACGTAGTCTAGTGGCCAGTTGTGTGTTAGCGTGGGGCTTTTGTCTA 679
Db 645 ACATGCAGATATATACAGCTCTCGTGGCAACATTTGGCTTCTCTGTGGTTATTTGCAA 586
Qy 680 GCAGTGAAGTGAACATTTGAGCAGTGAATGGAATACTACAAACATGGGAAGTATCCT 739
Db 585 GTGTGGAATGATGATCTTTAAGAGAGATGCAATCCAGTCTGGGAAGCTGTAT 526
Qy 740 ACATATGAACCTAGTGTGGACAGCTGTATCTGGCAGTTATTTCTCCATCGGTGGCACAG 799
Db 525 ATGTAATGACACTGCTGTGGACGCTATATCTTGGCAGTAGCATCAGTTGGAGTGTGG 466
Qy 800 GACTGATCTCGAGCTCTCTCTCATTTCTCAATGCAATGAAGCGTTTGGGACCTCCAG 859
Db 465 GATTGATCTTTGTGTGTCTATCGCTGTTTCAAAATGTGATAAGCACCCCTAGCTCTACCCA 406
Qy 860 TGGTTCCCTATCTTGGCTGTATCAATTTTCCATGACAAAATGAATGGGTTAAAGGTGATT 919
Db 405 TCATTCTCTGTTTGTCTGTGATTTTCTTTTCATGACAAAGATGGAGTAAAGATTTATAG 346
Qy 920 CTATGATCCTAGCTATTTGGGTTTCACTTCTATGTCTACCAACATATCTTGTATGACA 979
Db 345 CTATGCTGATGCCATTTGGGATTTATGTATATGGCCACCAATATATGTTGATGGCA 286
Qy 980 AAACTTGAAGAAAAT 996
Db 285 AGAAAGGTAGAAAGACT 269
```

RESULT 5
US-10-424-599-102519
; Sequence 102519, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B

```
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 102519  
; LENGTH: 2214  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63592C.1  
US-10-424-599-102519
```

Query Match 28.1%; Score 303.8; DB 16; Length 2214;
Best Local Similarity 60.2%; Pred. No. 2.5e-75;
Matches 540; Conservative 0; Mismatches 352; Indels 5; Gaps 2;

```
Qy 87 AGGAGTACTATGACAAACGAGGAGAAACAGTAAATGGCTAGCAACGGTAGTTCACATTGT 146
Db 476 AAGATTTTATATGATCAGGGTGAATAGTAAATGATGGCTACTCTAGTTCAAACTGC 535
Qy 147 TGGCTTTCTGTGCTACTTCCATATATATCTTGTGATTTTAAACACATGCAACACTGA 206
Db 536 TGGCTTCCCGATCTTGTTCATTCAT---TATTTACAAATTCCTTACCTCCAGAGGCTC 592
Qy 207 TAGAGATGGAAGAAAGAACCTCACCTAGGAACCGGTATTTGGTTTACGTAGTGTGGACT 266
Db 593 AACTTCTGCTTCACTCCCATCAAAATTAATCTTTGATATAT--TTGGTCTTTGGAGT 650
Qy 267 TCTTGAGGAGCAGATTGCTATCTGTACTCCATTTGGACTTCTTTACTTACCGCTTCTAC 326
Db 651 CTTAAATTTGCTGCTGACAAATATGATGTAATCCACTGGACTCTTATACCTCTCGGCTTCTAC 710
Qy 327 CTATTCCTGATCTGTGCTACTCAGTTAGCTTCAATGCTTTCTTCTCTTATTTCTTAA 386
Db 711 CTATTCCTGATTTGTGCTACACAGTTAGCTTTTAATGCAAGTTTCTCATATTTTATCAA 770
Qy 387 CTCAAAAACTTACCCTTATCTTTTAAATTTCTTTTCTTCTTAACTATATCTTCCAC 446
Db 771 TTCTCAAAAGTTCACTGCCCTTGATTAATACTCTACAGTGGTCTCACTTTTATCTGCTGC 830
Qy 447 CCTACTTTGCATTCATTAATGAGGAGACAGACTCCAAAAAGTTTACAAAAGAGAGTATGT 506
Db 831 ACTCCTTGTGTTAAACGAAGACACAGATGAACCATCTGGTTTCTCCAAGGAGAAATGACAT 890
Qy 507 CAAAGGTTTCATATGACACCGTTGCTGCTGCTGCTGTTATGTTAGTCTAGTCTTATCCCTACA 566
Db 891 TATTGGTTTCTATGTACCTTGGAGCTTCTGAGTGTACTCTCTTTTGTCTTTTCCCTCAT 950
Qy 567 ACAGTACGCTTTCTAAAAGTCTTAAAGAACAAAAATTTCTCAGAAAGTTATGATATGAT 626
Db 951 GCAGCTGACCTTTGAGAGAGTTCTGAAGAGAGGAAACATTTTCTGTGTTTGGAAATGCA 1010
Qy 627 AATCTAGTGAAGTCTAGTGCCAGTTGTGTAGCGTGGTGGGGCTTTTGTGCTAGCAGTGA 686
Db 1011 AATCTACACATCATTTCTGCTGCTGCTGCTCTGTCTATAGGCCCTATTTGCAAGTGGGA 1070
Qy 687 GTGMAAACCTTTGAGCAGTGAATGGAATACTACAAACATGGGAGGATATCTTACATATAT 746
Db 1071 ATGGCGTACTTTGCATGGAGAAATGGAGGGTTTTCAGAAAGGATATGTGCTTATGTTAT 1130
Qy 747 GAACCTAGTGTGGACAGCTGTTACCTGGCAGTTTATTTCTCCATCGGTGGCACAGACTGAT 806
Db 1131 GACTTTGGTTTGGACTTCAATAGCTGGCAGGATATGCTCTGTGTTGTGTTGCTGTTGAT 1190
Qy 807 CTTGAGACTCTCTCTCTATTTCTCAAAATGCAATGAGCGTTTGGGACTCCACAGTGGTCC 866
Db 1191 CTTCTAGTGTCTCTCTCTACTCCAATGTTATAAGCACAGTTTCTTTAGCCGCTAACTCC 1250
Qy 867 TATCTTGGCTGTAATCATTTTCCATGACAAAATGAATGGGTTAAAGGTGATTTCTATGAT 926
Db 1251 TATTGCTGCTGTATAGTTTTCATGATAAGATGAATGGGGTGAAGATAATTTCTATGCT 1310
Qy 927 CCTAGCTATTTGGGGTTTCACTTCTCTATGTCTACCAACAATATCTTGTATGACAAAAA 983
```

Db 1311 TTTGGCTCTATGGGTTTCCCTCTTATATTTATCAGAATTATCTTGATGATTCAAA 1367

RESULT 6
US-10-437-963-43860/c
; Sequence 43860, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 43860
; LENGTH: 2381
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2381)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_46977C.1
US-10-437-963-43860

Query Match 24.2%; Score 262; DB 17; Length 2381;
Best Local Similarity 59.1%; Pred. No. 2.1e-63;
Matches 448; Conservative 0; Mismatches 310; Indels 0; Gaps 0;

QY 235 AACCGTGATTTGGTTTACGTTAGCTGTGGACTTCTGTAGGAGCAGATTGCTATCTGTAC 294
Db 2039 AAGTCGCGGTGATATACATCGTCTGGGCTCATCATCGCCGCGACGACATGATGATAC 1980

QY 295 TCATTGGATCTTTTACTTACCCGTTTACCTATTCCTGATCTGTCGATCTCAGTTA 354
Db 1979 ACTGTGGGCTCAAGTACCTCCCGCTTCGACCTACTCGTCACTGTGCGCAGCAGCTC 1920

QY 355 GCGTCAATGCTTTCTTCTTATTTTCTTAACTCAGAAACTTACCCCTATCATTTTA 414
Db 1919 GCGTCAATGCTGCTTCTCATACGCTCAACTCCAGAGGTCACTCTCTGTATATTC 1860

QY 415 AATTCCTCTTTTCTCTTAACTATATCTCCACCTACTTGTGATTTCAATAATGAGGAGCA 474
Db 1859 AACTCGTGTGCTGCTCACCATGTCGCTTCGCTCATCGGAGTCAGCAAGAGTCTAG 1800

QY 475 GACTCCAAAAAGTTACAAAAGAGAGATATGTCAGAGTTTCATATGCAACCGTTGCTGG 534
Db 1799 GGGGTCAACGGGCTCTCGGAGGGAAGTATCTGCTCGTTCGTGCTGACGCTGGGGGG 1740

QY 535 TCTGCTGGTTATGGTCTACTCTTATCCCTACACAGCTAGCCTTCTTAAAGTCTTAAG 594
Db 1739 TGTGCACTACTCTGCTGATCTCGCGCTGATGCACTCACCTTCGAGACCATCATCAAG 1680

QY 595 AAGCAAAATTTCTCAGAAAGTTATGGATATGATAATCTAGCTGAGTCTAGTGCCAGTTGT 654
Db 1679 AAGCACACCTTCTCAGCCGTCTCAACATGCGATCTACCGCGCTGCTGGCGAGGCC 1620

QY 655 GTTAGCGTGGTGGGCTTTTCTAGCAGTGTAGTGAAACTTTTGAGCAGTGAATGGAT 714
Db 1619 GCGTGGTGGTGGGCTGTTTCGAGCGCGAGTGGAGGTCTGAGGGGGGAGATGAAC 1560

QY 715 AACTACAAACATGGGAGGATATCTTACATTATGAACCTAGTGTGAGCAGCTGTTACTGG 774
Db 1559 GCGTTCAGGTCGGGGCAGTTCTCCTACCTGATGAGCTGCTGTGGGCGGCCGCTGCTGG 1500

QY 775 CAGTTATTCTCCATCGGTGGCACAGGACTGATCTTCGAGCTCTCTCTCTCTATTTCTCAAA 834
Db 1499 CAGGTGGCCAAACATCGGGGTGCTCGGCTCATCTTCGAGGTGCTCGGCGCTCTTCTCCAAC 1440

QY 835 GCAATAAGCGTTTGGGACTCCAGTGGTTCCTATCTTGGCTGTGAATCAATTTTCCAGAC 894
Db 1439 GTGATCAGCAGCGTGTCACTCGCGGTCAATCCCGTCTTCGCGGTGCTGTGTTCCAGAC 1380

QY 895 AAAATGAATGGGTAAAGTGATTTCTATGATCCTAGCTATTTGGGGTTTCACTTCCTAT 954
Db 1379 AGGATGAACGGGTGAAGATCGTGCCATGCTGATTCGAATTTGGGGATTTATTCGTAT 1320

QY 955 GTCTACCAACAATATCTTGATGACAAAAAATTGAAGAA 992
Db 1319 CTGTTTCAGCACTATCTAGATGGCAAGAAAGCAAGAA 1282

RESULT 7
US-10-416-898-9
; Sequence 9, Application US/10416898
; Publication No. US20040172670A1
; GENERAL INFORMATION:
; APPLICANT: Yale University
; APPLICANT: Walker, Elisabeth
; APPLICANT: Dellaporta, Stephen
; TITLE OF INVENTION: MAIZE YELLOW STRIPE1 AND RELATED GENES
; FILE REFERENCE: 44574-5106-US
; CURRENT APPLICATION NUMBER: US/10/416,898
; CURRENT FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: PCT/US01/43101
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 60/249,222
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 83698
; TYPE: DNA
; ORGANISM: Zea mays
US-10-416-898-9

Query Match 21.7%; Score 234.2; DB 17; Length 83698;
Best Local Similarity 54.3%; Pred. No. 1.3e-54;
Matches 495; Conservative 0; Mismatches 413; Indels 3; Gaps 1;

QY 92 TATACTATGACACGAGGAAACAGTAATGGCTAGCAACGGTAGTTCAACTGTGTGGCT 151
Db 67301 TTTATTACAATGAAGGTGGAACAGTAATGGATCTCTACTCTTGTTCAAACTGGTGGCT 67360

QY 152 TTCTGTGCTACTTCCATATTTATTTGTCATTTTAAACACATGCAACAACGATAGAG 211
Db 67361 TTCCATTTCTTTATCTCCCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 67417

QY 212 ATGAAAAAAGAACCTCACCTAGGAACCGTGTATTTGGTTTAGCTAGTGTCTGGACTTCTTG 271
Db 67418 CT 67477

QY 272 TAGGAGCAGATGCTATCTGTACTCCATGAGCTTCTTTACTTACCCGTTTCTACCTATT 331
Db 67478 TTGGTTTAGATAAATTTTTTATCTCTGTGGACTTGTGTATCTCTCTCTCTCTCTCTCT 67537

QY 332 CCCTGATCTGTGCATCTCAGTTAGCTTCAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 391
Db 67538 CGATTCTATGTGCTTCAAGTTAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 67597

QY 392 AAAAATCTTACCCCTATCATTTTAAATCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCT 451
Db 67598 AGAAATCACTTGTGTGATTTCTCTCAGTGTGTTTCTCTCTCTCTCTCTCTCTCTCTCT 67657

QY 452 TTGCAATCAATAATGAGGAGAGCAGACTCCAAGATTTACAAAAGAGAGATATGTCAAAG 511
Db 67658 TTTCTCTTGAGGATGATTTCAAAATAGCCCATCAGGAGATCTTAAAGTGGAGTACTTGAATG 67717

QY 512 GTTTCATATGACCGTCTGGTCTGCTGGTTATGGTCTAGTCTTATCCCTCAACAGC 571
Db 67718 GGTGTTCTGTGACGTTTGTGCTCTCTATCTCTCTTCTGAGCTCTCTTATGACGT 67777
QY 572 TAGCCTTTCTAAAGTCTCTAAAGAGCAAAATTTCTCAGAAAGTTATGATATGATAATCT 631
Db 67778 TTTCTTTTCGAGAGGTTCTCAGAGTGAGACTCTCTCTATGGTCTCGAGATGCATACT 67837
QY 632 AGCTGAGTCTAGTGCCAGTTGTGTTAGCGTGGTGGGCTTTTGTCTAGCAGTGAAGTGA 691
Db 67838 ATACGCTGCTGTGGCTTCTTGTGTAGCGGTTATCGGATTTGTGCAAGCGGGAAATGA 67897
QY 692 AAATCTTCAGCAGTGAATGATGATCAACAACTGGAAGGTATCTTACATTTATGACCT 751
Db 67898 TGTGTTGAGTGTGGAGATGGAAGGTTTTCAGGAAGGTCAAGTCAATTTATGTTTGTGACTT 67957
QY 752 TAGTGTGACAGCTGTTACCTGGCAGTTTATTTCTCCATCGGTGGCAGCAGTGTATCTTCG 811
Db 67958 TGGTCGGGGCAGCGGTTTCGTGTCAATTTGGGTTGTGTAGGAGCCGTGTCTTATTTTC 68017
QY 812 AGCTCTCTCTATCTCAATGCAATGAAGGTTTGGGACTCCCGAGTGGTCTCTATCT 871
Db 68018 TGGTGTCTGCTGTGTTTCAAACTTATAGTACGCTCTCACTCATTTGTTACGCTCTCTCG 68077
QY 872 TGGCTGTAAATCATTTTCCATGACAAATGAATGGTTAAAGGTGATTTCTATGATCCTAG 931
Db 68078 CGGCCAATTCGGGTGTTCATGACAAAGCTGACTGAGGTTAAGATGGTGGAGTCCCATCG 68137
QY 932 CTATTGGGGTTTCACTTCTCTATGCTACCAACAATATCTTGATGACAAAACTTGGAAG 991
Db 68138 CCTTCACAGGATTCAGTTTATATCTACCAAGACTATCTTGATGACTTGAAGTACAAA 68197
QY 992 AAAATCATGAA 1002
Db 68198 GAGCAGAGAA 68208

RESULT 8

US-10-424-599-75161
; Sequence 75161, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 75161
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_38884C.1
US-10-424-599-75161

Query Match 20.7%; Score 223.4; DB 16; Length 997;
Best Local Similarity 58.4%; Pred. No. 1.2e-52; Mismatches 30; Indels 30; Gaps 3;
Matches 464; Conservative 0;

QY 87 AGGAGTACTATGACAAAGGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCAACTTGT 146
Db 202 AAGATTATCTATGACATGGTGGTAAATGCAATGGATGGCAACATTTGTTCAATCAGC 261
QY 147 TGGCTTTCTGTGCTACTTCCA---TATTATATCTTGTTCATTTAAACACATGCAACAC 203
Db 262 TGGATTCCGTGTACTTCTTCCACTCTCTTTTACTTCCCAAGACAAACACATGCCAAGTT 321

QY 204 TGATAGAGATGG-----AAAAAGAACTCTCACCTAGGAACCG 239
Db 322 TAATACAAATCCAGCAATAATGATATTCTCAAAACAAACCAAAATTTCTCCACCT 381
QY 240 TGATATTGTTTACGTAGTGTGGACTTCTGTAGGAGCAGATTGCTATCTGTACTCCAT 299
Db 382 TGTCTTCTCTATCTAGCTTTTGGCCTAAATCTCACAGGGGCAACCTTGATGTATTCTTA 441
QY 300 TGGACTTCTTACTTACCCGTTTCTACCTATTTCCCTGATCTGTGCATCTCAGTTAGCTT 359
Db 442 TGGACTTTTATATCTCCCTCTTTCCACCTATTTCTACTATGTGCGCAACCAATTTAGTCTT 501
QY 360 CAATGCTTTTCTCTTATTTTCTTAACTCAAAAACTTTACCCCTATCAATTTTAAATTC 419
Db 502 CAAGCGGTGTTCTTTTCTCTCAATTTCCGAAATTTACAGCATTCATATCAATTC 561
QY 420 TCTTTTCTCTTAACTATATCTTCCACCTTCTGCAATTCATTAATGA---GGAGACAGA 476
Db 562 TGTAGTCTCTTACCATAATCAGCTTCTCTGCTTGCATCAACTCCGATTCGACAGGAG 621
QY 477 CTCCAAAAAGTTACAAAAGGAGATGTCMAAGGTTTTCATATGACACCGTTGCTGCGTC 536
Db 622 CTCAACCGGCTTTTCCAGAGAAAGCATGTAATTTGGATTTCTTGCACCAATTTGGTGCATC 681
QY 537 TGCTGGTTTATGGTCTAGTCTTATCCCTCAACACAGCTAGCCTTTCTAAAAAGTCTTAAAGAA 596
Db 682 AGCTACATTTCTTTGTACCTCTCTTGTGAGCTTTCTTTTCAGAAAGTTATAAGAG 741
QY 597 GCAAAATTTCTCAGAAAGTTATGGATATGATAATCTACGTGAGTCTAGTGGCAGTTGTGT 656
Db 742 AGAAACCTTTTCTGCTGTGTGGACATGCAATTTTACCATCTTCTTGTACATGTGC 801
QY 657 TAGCGTGTGGGCTTTTGTCTAGCAGTGTGAGTGGAAAACCTTTGAGCAGTGAATGGATAA 716
Db 802 TTGTGTAGTAGGGTTGTTGCAAGTGGAGAGTGGAAAAGTTTGAACAAATGAGATGAAGG 861
QY 717 CTACAAACATGGGAAGGTATCTACATTTATGAACCTAGTGTGGACAGCTGTACTCGCA 776
Db 862 ATATGATAAGGAAGTGTGCCGTTGTAATGACTGTACTATAGATTCTGTGACATGGTA 921
QY 777 GTTATTCTCCATCGGTGGCAGCAGGACTGATCTTGGAGCTCTCTCTCTATTTCTCAAATGC 836
Db 922 TATATGCTCACTAGGATGCTAGGGGGAGCTTTGGAGGAGTGTGCAATTTGTTGCGGATGT 981
QY 837 AATAAGCGTTTGGG 851
Db 982 GAGAAGTACCTGGGG 996

RESULT 9

US-10-767-701-9000
; Sequence 9000, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 9000
; LENGTH: 728
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS69833_1
US-10-767-701-9000

Query Match 16.6%; Score 179.2; DB 17; Length 728;
Best Local Similarity 61.2%; Pred. No. 3.9e-40;

Matches 289; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 532 GCCTCTGCTGGTTATGCTAGTCTTATCCCTACACAGCTAGCCTTTCTAAAGTCCTAA 591
Db 4 GCCTCAGCCACATACCTCACTAATCTCTCTAATGCAAGTCACATTTGAGAGGTTATT 63
QY 592 AGAAGCAAAATTTCTCAGAAAGTTATGGATATGATATCTACGTGAGTCTAGTGGCAGT 651
Db 64 AAGAAGGAAACGTTCTCAGTCGTGTTGAATATGCAGATATATACGACACTAGTGGCAACA 123
QY 652 TGTGTTAGCTGCTGGGGCTTTTGTCTAGCAGTGAGTGGAACCTTTGACGAGTGAATG 711
Db 124 ATAGCTTCTCTGTTGGGTTATTGCAAGCGGTGAGTGAAGACTTTAGAGGAGAGATG 183
QY 712 GATAACTACAAACATGGGAAGGTATCTCATATATGAACCTAGTGTGGACAGCTGTACC 771
Db 184 CATGCCCTTCAGCTCAGGAGGTTGTCTATATGATGACACTTCTATGGACTGCTGTATCT 243
QY 772 TGGCAGTTATTCCTCAATCGGTGGACAGGACTGATCTTCGAGCTCTCCTCTCTATTCTCA 831
Db 244 TGGCAGATAGCTTCTGTAGAGTGGTGGGTTGATCTTTGTGTGTCATCACTCTTTTCA 303
QY 832 AATGCAATAGCGTTTGGGACCTCCAGTGGTTCCTATCTTGGCTGTAATCAATTTCCAT 891
Db 304 AATGTGATAGCACACTAGCTCTACCACATCAITCCAGTTTTTGTGTGATTTTCTTCCAC 363
QY 892 GACAAATGAATGGGTTAAAGGTGATTTCTATGATCTAGCTATTTGGGGTTTCACTTCC 951
Db 364 GACAAGTGAATGGAGTGAAGATTTATGTTATGTTGATGGCCATCTGGGATTTGTTTCA 423
QY 952 TATGCTACCAACAATATCTTGTATGACAAAACTTGAAGAAAATCATGAAA 1003
Db 424 TATGGATACCAATATATATATAGTACAGAAAGCTTAGGAAGACTTCAGTAA 475

RESULT 10
US-10-424-599-114931
; Sequence 114931, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 114931
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_74793C.1
US-10-424-599-114931

Query Match 13.8%; Score 149; DB 16; Length 522;
Best Local Similarity 56.8%; Pred. No. 1.3e-31;
Matches 295; Conservative 0; Mismatches 225; Indels 1; Gaps 1;

QY 353 TAGCCCTCAATGCTTCTCTTATTTTCTTAACCTCAAAAACCTTACCCCTATCATTT 412
Db 1 TAAATTTCAATGAGTGCTCACTTTCTTCATCAATTTCCAAAAGTTCACTGCATGATCT 60
QY 413 TAAATTTCTTTTCTTCTTAACCTATATCTTCCACCCTACTTGCATTTCAATAATAGGAGA 472
Db 61 TAAACTTATATCGCTTACCATATCATGTTACCTGATTTGCACCTCAATCACTGATCTG 120
QY 473 CAGACTCCAAAGTTACAAAGGAGATGTCTAAAGGTTTCATATGCAACGTTGTG 532
Db 121 AGSAAACAAAGAACCTTTCCAAACAGAGCAAAATATTTGGGTTCTTCTGTGCCCTAGGTG 180

QY 533 CGTCTGCTGGTTATGCTAGTCTTATCCCTACACAGCTAGCCTTTCTAAAGTCCTAA 592
Db 181 CATCTGCCATATTTGCATTCATCTCTCTTATGCAATTTATTTTGAGAAAATTTATA 240
QY 593 AGAAGCAAAATTTCTCAGAAAGTTATGGATATGATATCTACGTGAGTCTAGTGGCAGTT 652
Db 241 AGACTGAAACCTTTTCTACTGTATTAAAGCATGATTTTTTACCCAAATGATCGTGGGTACAA 300
QY 653 GTGTTAGCTGCTGGGGCTTTTGTCTAGCAGTGAGTGGGAACCTTTGACAGTGAATGG 712
Db 301 TTGGTGGCCTTGTGGGATTTGTTAAAGTGGAGATTGGGAACTATGGGAATGGAGATGA 360
QY 713 ATAACCTACAAACATGGGAAGGTATCTCATATATGAACCTAGTGTGGACAGCTGTACCCT 772
Db 361 AGGAATTCGAAAACGGTAGTGTCTATATGATGACTCTGTTTGTACTTCTGTGACAT 420
QY 773 GCGAGTTATTTCTCAATCGGTGGCAGGACTGATCTTCGAGCTCTCCTCTCTATTCTCA 831
Db 421 GGCAGATAGTGTGTGTTGATGCTGGGGTTGATTTTTTGAGGTATCTTCATTACCTCAG 480
QY 832 AATGCAATAGCGTTTGGGACTCCGAGTGGTTCCTATCTT 872
Db 481 CCTGCATAAGCAATCTGGAATTGAAACATAGCTCTCTATCTCT 521

RESULT 11
US-10-425-114-12485
; Sequence 12485, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 12485
; LENGTH: 1386
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 701179255_FLI
US-10-425-114-12485

Query Match 13.1%; Score 141.6; DB 16; Length 1386;
Best Local Similarity 50.5%; Pred. No. 2.9e-29;
Matches 370; Conservative 0; Mismatches 359; Indels 3; Gaps 1;

QY 248 TTTACGTAGTGTCTGGACTTCTTTAGGAGCAGATTTGCTATCTGTACTCCTTCAATGGACTTC 307
Db 411 TCTACGCGGCATCGGGGTCTATGATCGGCTTCGACAACTCATGTACTCTGTACGGCTGC 470
QY 308 TTTACTTACCGGTTTCTACCTATTTCCCTGATCTGTGCACTCTCAGTTAGCCTTTCAATGCTT 367
Db 471 AGTACTCGCGGTGCCACCTTCTCGTCTGGCGCGCAGCAGCTGGGCTTCAACGCCA 530
QY 368 TCTTCTCTTATTTCTTAACCTACAAAACCTTACCCCTATCATTTTAATTTCTCTTTTCC 427
Db 531 TCACCTCGCGCTCATCAACGCGAGCGGTTTACGCGGCTGATCGCCAACTCCGTGGTGG 590
QY 428 TCTTAACCTATATTTTCCACCCCTACTTTGATTTCAATTAATGAGGAGACAGA---CTCCACAA 484
Db 591 TGCTCACCTTCTCGCGCGGATCTCTCGCGGTGGATTCCTCTCCGACGAGACCTCCAGA 650
QY 485 AAGTTACAAAAGGAGAGTATGTCAAAAGGTTTCATATGCAACCGTTTCTCGCTCTGCTGTT 544

Db 651 ACGTCGCGCGGCAAGTACCACCTGGGGTTTCTGACTCGTCTGGCGCCCTCGGCGTGT 710
Qy 545 ATGGTCTAGTCTTATCCCTACAAAGCTAGAGCTTTCTAAAGTCTCTAAAGAACAAATTT 604
Db 711 TCGCGCTCAATTTTGTCCCTCTTCGAGCTTCACCTTCGAGAAGGTGGTCAGGGTGGGACGG 770
Qy 605 TCTCAGAAGTTATGATATGATAATCTACGTGAGTCTAGTGGCCAGTTGTGTAGCGTGG 664
Db 771 CGCGTGGGTGCTGCGGATGAGATGACACCACTGGTGGCGTTCGCGGTGTCGCGTGG 830
Qy 665 TGGGCTTTTGTAGCAGTGGAGTGGAAACCTTTGAGCAGTGAATGGATAAATACAAAC 724
Db 831 CGGGCTGCTCGCGTGGGGGACTGGCGGACGATCCGGGGGAGATGGCGTCTCAAAG 890
Qy 725 ATGGGAAGGTATCTACATATATGAACCTAGTGTGGACAGCTGTTACTTGGCAGTATTTCT 784
Db 891 ACGGAGACGAGGTACGTGCTGAGCGTGTGGGACGCGCGGTGTCGTGGCAGGACGG 950
Qy 785 CCATCGGTGGCACAGGACTGATCTTCGAGCTCTCTCTATTTCTCAAAATGCAATAAGCG 844
Db 951 CGTGGGCTTGGTGGCGGTGATCATGAGGCTGTGCTGCTGTTCCGGAACGTGACGTGCA 1010
Qy 845 TTTTGGGATCCCGATGGTTCCTATCTTGGCTGTAATCAATTTTCCATGACAAAAATGAATG 904
Db 1011 CGTGGCGCTGCGCTGTGCGCAGTGTGCGGTGGCGCTGTTGGGGACAGGATGACCG 1070
Qy 905 GGTAAAGGTGATTTCTATGATCTCTAGTATTTGGGTTTCACTTCTCTATGCTTACCAAC 964
Db 1071 GTATAAAGATCGTGGCCATGCTCATGGCGGTATGGGGGTTCTCTCTACATGTACCAGC 1130
Qy 965 AATATCTTGATG 976
Db 1131 ACTATCATCGACG 1142

RESULT 12
US-09-770-152-44
; Sequence 44, Application US/09770152
; Publication No. US20020040489A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickter, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2025US (PARA-014PRV)
; CURRENT APPLICATION NUMBER: US/09/770,152
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,503
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-152-44

Query Match 12.8%; Score 138.6; DB 9; Length 592;
Best Local Similarity 60.2%; Pred. No. 1.3e-28;
Matches 250; Conservative 0; Mismatches 159; Indels 6; Gaps 1;
Qy 59 TCACCTCTAAACACATACAAACGGTGGCTTCAGGAGTATATGACAAAGGAGGAACAGTA 118
Db 176 TCTGCCAACACATCTGTACAGTTCTGGGTAGACTGTACTATGAAATGGAGGGAAGCA 235
Qy 119 AATGGCTAGCAAGGTAGTTCAACTTGTGTGGCTTTCTGTGCTACTTCCATATATATCT 178
Db 236 CATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGATTTCTTCCGCTTCT 295
Qy 179 TGTCAATTA-----AAACACATGCGACAACTGATGAGATGGAAGAAACCTCACTA 232
Db 296 TTTTCGAAATCAGCAACCCAAATCAACAGATCAAAATTTTCAGTCAGTCCCTTCTTCA 355
Qy 233 GGAACCGTGTATTGGTTTACGTAGTGTCTGGACTTCTTGTAGGAGCAGATTTGCTATCTGT 292
Db 356 CCACCTTGCATCGGTTTACTTGTGCATGAGCTGTAGTGTCTGCTTATGCTTATTTGT 415
Qy 293 ACTCATTTGACTTCTTTTACTTACCGGTTTCTACCTATTTCCCTGATCTGTGCACTCAGT 352
Db 416 CTGCAGTTGGGTGCTTTACTTACCAGTCTCTACTTTCTCCCTCATCTTGGCCTCACAGT 475
Qy 353 TAGCCTTCAATGCTTTTCTCTTATTTTCTTAACACAAAAAATTTACCCCTATCATTT 412
Db 476 TGGCTTTTCACTGCTTTTCTCATATTTCTTAACTCGAAAAAGTTCACTCCTTTGATAG 535
Qy 413 TAAATCTCTTTTCTCTTAACATATATCTTCCACCCCTACTTGCATTTCAATAATGA 467
Db 536 TCAATCTTTTCTCTTACGGTTTCTCTGCGCTCCTCTGCTGGTCAACACTGA 590

RESULT 13
US-10-425-114-28473
; Sequence 28473, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 28473
; LENGTH: 1779
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4757-034-H3_FLI
US-10-425-114-28473

Query Match 12.8%; Score 138.4; DB 16; Length 1779;
Best Local Similarity 50.3%; Pred. No. 2.7e-28;
Matches 368; Conservative 0; Mismatches 361; Indels 3; Gaps 1;
Qy 248 TTTACGTAGTGGTGGACTTCTTGTAGGAGAGATTGCTATCTGTACTTCCATTGGACTTC 307
Db 775 TCTACGGGGCATCGGGGTCTATGATCGGCTTCGACAACTCTATGTACTCTGACGCGTGC 834
Qy 308 TTTACTTACCGTTTCTTACCTATTTCCCTGATCTGTGATCTCAGTTAGCCTTCAATCCTT 367
Db 835 AGTACCTGCGGGTGTCCACCTTCTCGCTCGTGGCGGACGACGCTGGGTCTTCAACGCCA 894
Qy 368 TCTTCTCTTATTTTCTTAACCTCAAAAAAATTTACCCCTATCATTTTAAATTTCTTTTCC 427

895	TCACTCGCGCTCATCAACGGCGACGGTTCACGGCGCTGATCGCCAACTCCGTGGTCG	954
428	TCTTAACATATATCTTCCACCCTACTTGCATTCAATAATGAGGAGACAGA---CTCCACAA	484
955	TGCTCACCTTCTCGCGCGGATCCTCGGCGTCGGATCCTCTCCGACGAGACCTCCAGCA	1014
485	AAGTTACAAAAGGAGAGATGTCAAAGTTTCATATGACACGTTGCTCGCTCTGCTGTT	544
1015	ACGTGCGCGCGGCAAGTACCGCTCGGGGTTTCGCTCTGCTGCGCCCTCGCGGTGT	1074
545	ATGGTCTAGTCTTATCCCTCAACAGCTAGCTTCTTAAAGTCTCTAAAGAAGCAAAATT	604
1075	TCGCGCTATTTTGCTCTTCGAGCTCACCTTCGAGAGGTGGTCAAGGTGCGACGG	1134
605	TCTCAGAAGTTATGGATATGATAATCTACGTGAGTCTAGTGGCCAGTTGTGTAGCGGTG	664
1135	CGCGTGGGTGTCGGATGCAGATGCACACAACCTGGTGGCGTCGCGCTGTCGCTGG	1194
665	TGGGGCTTTTGGCTAGCAGTCAGTGGAAAACCTTTGAGCAGTGAATGATACTACAAC	724
1195	CGGGCTGCTCGCTCGGGGACATGGCGGACGATCCCGGGGGAGATGGCGTCGTTCAAG	1254
725	ATGGGAAGGTATCTACATATGAACCTAGTGTGGACAGCTGTACCTGGCAGTTATTCT	784
1255	ACGGGAGACGAGGTACGTGCTGACGTGTGGGACCGCGGTGCTGTGGCAGGACGGG	1314
785	CCATCGTGGCACAGGACTGATCTTCGAGCTCTCCTCTATTTCTCAAAATCAATAAGC	844
1315	CCGTGGGCTTGGTCGGCTGATCATGAGGTGTCTGCTGCTTCGGAACGTGACGTGCA	1374
845	TTTTGGGACTCCAGTGGTTCCATCTTGGCTGTAAATCATTTTCATGACAAAATGAATG	904
1375	CGCTGGCGCTGCCGCTGGTGCAGGTGTTGCGGTGGCGCTGTTCCGGGACAGGATGACG	1434
905	GTTTAAAGGTGATTCTATGATCCTAGCTATTTGGGGTTTCACATTCTCTATGCTACCAAC	964
1435	GTATAAGATCGTGGCCATGCTCATGGCGGTACGGGGTTCCTCTCTCATGTACCGC	1494
965	AATATCTTGATG	976
1495	ACTACATCGACG	1506

RESUM.T 14

```

US-10-424-599-30274
; Sequence 30274, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 30274
; LENGTH: 1233
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12733C.1
US-10-424-599-30274

```

	Query Match	12.2%	Score 131.8	DB 16	Length 1233
	Best Local Similarity	47.3%	Pred. No. 1.7e-26		
	Matches 431	Conservative 0	Mismatches 477	Indels 3	Gaps 1
Qy	92	TATACTATGCAACGGAGGAAACAGTAAATGGCTTAGCAACGGTAGTTCAACTGTTGGCT	151		
Db	8	TCCTACTTCATCCACGGTGGCAACGAATCTGGCTCTCTAGCTTCCTTGAACATGCACGCA	67		

QY	152	TTCTGTGCTACTTCCGATATATATCTGTGCAATTTAAACAACATGCACCAACTGATAGAG	211
Db	68	ATTCCGGCTCGAGTCTCTCCCATTAATCATTTCTTACATTCATAAAGCAGCTCGTCATCGCT	127
QY	212	ATGGAAAAAGAACCTCACCTAGGAACCGGTAT--TGGTTTACGTPAGTGTGGAATTTC	268
Db	128	CTTCGGCCCTAATCTATCAAGCCCTCTCTTCTCGCCTCGGCCCTCATCGGACTCG	187
QY	269	TTGTAGGACGAGTTGCTATCTGTACTCCATTTGGACTTCTTTACTTACCCGTTTCTTACT	328
Db	188	TCACCGGGCTCGACGACTACTCTACGCTGCGGGCTGCTCGCTTCCGGTCTCCACTT	247
QY	329	ATTCCCTGATCTGTCATCTCAGTCTAGCTCAATGCTTCTTCTCTTATTTCTTTAACT	388
Db	248	TCTCTTAATCAAGCCTCCCACTCGCCTTCAACCGCGTCTTCGCTTCTCTCTCGTCC	307
QY	389	CACAAAAACTTACCCCTATCATTTTTAAATTCCTTTTCTCTTAACTATATCTTCCACCC	448
Db	308	GCCACAGGTTCAAGCCCTACTCCGTCAACTCGTCTGCTCTTCAACCGTCGCGCTGTGG	367
QY	449	TACTTGCATTCATTAATGAGGAGACAGACTCCAAAAAGTTTACAAAAGGAGAGTATGTCA	508
Db	368	TTCTGGCTCTCGGTTCTCAGCGGGACCGCCCGCGGTGAGTCGAGTCGCCAGTACGTGA	427
QY	509	AAGGTTTCATATGCAACCGTTGCTGCGTCTGCTGGTTTATGGTCTAGTCTTATCCCTACAAC	568
Db	428	TGGTTTTGTTATGATCTTGGCGCTGCGGCTGTATGGGTTTGGTTTTGCCGTTGATGG	487
QY	569	AGCTAGCCTTTCTAAAAGTCTTAAAGAGCAAAATTTCTCAGAAGTTATGGATATGATAA	628
Db	488	AGTTGGTGTACAAAAGAGCAGGACGATATCACGTACTCTCTGGTCATGGAGATTCAAG	547
QY	629	TCTACGTGAGTCTAGTGGCCAGTTGTGTAGCGTGGTGGGGCTTTTGTCTACGACGTGAGT	688
Db	548	TGTCTGTGCTTCTTGTACCTTAATTCGACCGCTGGGAATGATATCAATAATGACT	607
QY	689	GGAAACCTTTGAGCAGTGAATGGAATACTACAAACATCGGAAGTATCCTACATTTAGA	748
Db	608	TCAAGGTATTTCCGCGAAGCAAGAGATTTTAAGCTTTGGGGAACAAAGTACTACGTTG	667
QY	749	ACCTAGTGTGACAGCTGTTACTGGCAGTTTATCTCCATCGGTGGCAGAGACTGATCT	808
Db	668	TGTTGGTGTGAGTGCAATAATATGGCAGTTTTTCTTCTGGAGCAATAGGGGTATCT	727
QY	809	TCGAGCTCTCTCTATTTCTCAAAATGCAATAAGCGTTTTTGGGACTCCACAGTGGTCCCTA	868
Db	728	TTTGTGCTCTGTTTGTGTGCGGTATTAATTTGCTGCTTTCTTCCAGTGAACGGAAG	787
QY	869	TCTTGGCTGTAAATCAATTTTCCATGACAAAAATGAATGGGTTAAAGTGATTTCTATGATCC	928
Db	788	TTTTGGCTGTATTGTATACAAAGAGAGCTTTTCATGCAGAAAAAGGGGTGCTTTGGTGC	847
QY	929	TAGCTATTTGGGGTTTCACTTCTTATGTCTACCAACATATCTTGTATCAGAAAACTTGA	988
Db	848	TCTCTTTTGGGGTTTGTGTCTTATTTCTATGGAGATTAACACAGACGGGAAAAAGA	907
QY	989	AGAAAAATCAT	999
Db	908	ACAAGAATCGT	918

RESULT 15

US-101-425-114-10516
; Sequence 10516, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei

GenCore version S.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:48:03 ; Search time 99.7262 Seconds
(without alignments)
7704.726 Million cell updates/sec

Title: US-09-913-767-6
Perfect score: 1081
Sequence: 1 caaatccaacagttcaagat.....ataagctgattttgaaag 1081

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PCITUS COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57.8	5.3	7218	1	US-08-232-463-14
2	46.8	4.3	1141	4	US-09-806-708B-22
3	42.2	3.9	832	4	US-09-621-976-2813
4	41.2	3.8	1141	4	US-09-806-708B-22
5	40.6	3.8	72604	3	US-09-268-992-7
6	40.6	3.8	72604	3	US-09-657-474-7
7	40.4	3.7	1055	4	US-09-806-708B-23
8	39.2	3.6	474	4	US-09-621-976-18033
9	38.6	3.6	2674	3	US-09-019-095A-1
10	37.8	3.5	1664976	4	US-08-916-421B-1
11	37.8	3.5	1664976	4	US-09-692-570-1
12	36.8	3.4	248	3	US-09-007-005-32
13	36.8	3.4	248	3	US-09-244-796-32
14	36.8	3.4	277	3	US-09-007-005-3
15	36.8	3.4	277	3	US-09-244-796-3
16	36.8	3.4	8093	4	US-10-204-708-31
17	36.6	3.4	1536	4	US-09-107-532A-2458
18	36.2	3.3	4726	2	US-08-633-770A-11
19	35.8	3.3	70000	4	US-09-851-896-3
20	35.6	3.3	24358	4	US-09-392-812A-1
21	35.4	3.3	705	4	US-09-270-767-5061
22	35.4	3.3	705	4	US-09-270-767-20343
23	35.4	3.3	909	4	US-09-710-279-3061
24	35.4	3.3	948	3	US-09-134-001C-1684
25	35.4	3.3	3554	4	US-09-710-279-4291
26	35	3.2	5217	4	US-09-100-703A-25
27	35	3.2	5230	4	US-09-100-703A-26

C 28	35	3.2	5230	4	US-09-838-718A-5	Sequence 5, Appli
C 29	35	3.2	5231	4	US-09-100-703A-27	Sequence 7, Appli
C 30	35	3.2	5770	4	US-09-838-718A-7	Sequence 7, Appli
C 31	35	3.2	5870	4	US-09-838-718A-8	Sequence 8, Appli
C 32	35	3.2	5906	4	US-09-838-718A-6	Sequence 6, Appli
C 33	34.8	3.2	548	4	US-09-621-976-894	Sequence 894, App
C 34	34.8	3.2	696	3	US-09-461-697-193	Sequence 193, App
C 35	34.8	3.2	699	3	US-09-461-697-191	Sequence 191, App
C 36	34.8	3.2	717	3	US-09-461-697-189	Sequence 189, App
C 37	34.8	3.2	774	3	US-09-461-697-187	Sequence 187, App
C 38	34.8	3.2	819	3	US-09-461-697-185	Sequence 185, App
C 39	34.8	3.2	1669	3	US-09-461-697-184	Sequence 184, App
C 40	34.6	3.2	832	4	US-09-621-976-2813	Sequence 2813, Ap
C 41	34.6	3.2	2319	4	US-09-561-077C-15	Sequence 15, Appl
C 42	34.6	3.2	2319	4	US-09-221-014-15	Sequence 15, Appl
C 43	34.6	3.2	3551	4	US-09-561-077C-16	Sequence 16, Appl
C 44	34.6	3.2	3551	4	US-09-221-014-16	Sequence 16, Appl
C 45	34.6	3.2	7113	4	US-09-561-077C-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEPLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 5.3%; Score 57.8; DB 1; Length 7218;

```

Best Local Similarity   8.5%; Pred. No. 4.9e-07;
Matches 17; Conservative 125; Mismatches 57; Indels 0; Gaps 0;

QY      261 TGGACTTCTTGTAGGACAGATTGCTATCTGTACTTCATTTGGACTTCTTTTACTTACC CGT 320
        ||| | | | | | | | | | | | | | | | : : : : : : : : : : : : : : :
Db      1039 TTGGCTGCAGGTGCAGGAGCTTGGCATYYYYYYYYYYYYYYYYYYYYYYYYYYY 1098

QY      321 TTCTACCTATTCCCTGATCTGTGCGATCTCAGTTAGCCTCAATGCTTCTTCTCTCTATT 380
        ||| | | | | | | | | | | | | | | | : : : : : : : : : : : : : : :
Db      1099 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1158

QY      381 TCCTAACTCACAAAACCTTACCCTATCATTTTAAATTCCTCTTTTCTCTTAATAATC 440
        ||| | | | | | | | | | | | | | | | : : : : : : : : : : : : : : :
Db      1159 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1218

QY      441 TTCACCCCTACTTGCATTC 459
        ||| | | | | | | | | | | | | | | | : : : : : : : : : : : : : : :
Db      1219 YYYYYYYYYYYYYYYYYYYY 1237

RESULT 2
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIORITY FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIORITY FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAE1 promoter
; US-09-806-708B-22

```

Db	709	WKAYYAHATNNWGWNNNTDARETNNNTVNRERW-----MINTKTRWISTTRR	662
Qy	531	TGGCTCTGCTGGTATATGGCTTAGCTTATCCCTACAAACAGCTAGCCCTTTCTAAAAAGTCCT	590
Db	661	HHVTGATNN	602
Qy	591	AAAGAAACAAAAATTTCTCAGAAAGTTATGGATATGATAATCTACGTGAGCTTAGTGCCAG	650
Db	601	VWADSWVWVYANWVRCDVTYTRNNTYCKSYAHSYVWYSSNAM-WYRRYSARNWSSMARW	543
Qy	651	TTGTGTTAGCGTGGTGGGGCTTTTGGCTAGCAGTGGTGGAAAACCTTTGAGCAGTGAAT	710
Db	542	TTTRNNWMSGBVRMRWAGTNMWRHNNNNNTDTRYWWKWARBTITTVYDSMCKNAKSMWR	483
Qy	711	GGATAACTACAAACATGGGAAGGTATCCTACATATATGAACCTAGTGTGGACAGCTGTTC	770
Db	482	GNWRAMKMWAAANDAGANDHWYWMGNNTMMRRAMKMMNAWCRAYCCNNNNNRAC	423
Qy	771	CTGGCAGTTATCTCCATCGGTGGCAGCAGCAGTATCTTCAGACTCTCCCTCTCTATTCCT	830
Db	422	VHHKHQWRWTWKYMKWKAACNNNNBRAMVYMRVAMMYSRDTTNTDWMWMTSDWBWHWTV	363
Qy	831	AAATGCATTAAGCGTTTGGGACTCCAGCTGGTTCCTATCTTGGCTGTATCATTTTCCA	890
Db	362	DYTMRAWNNNNNNNNNRBCKTTSGWMMMDHMHNTCTYGNNTWGSAYBMAAMSWAAGASN	303
Qy	891	TGACAAATCAATGGGGTAAAGGTGATTTCTATGATCCTAGCTATTTGGGGTTCACCTC	950
Db	302	BVTYNWCWRYMGKTYNN	243
Qy	951	CTATGCTCACAACAATATCTTGATGACAAAACTTTGAAGAAAAATCATGAAATCAAC	1010
Db	242	BWYBYMTGKHBBWRRRABHRSWNNWVVKCRNKYVMVSWHYHAMRYBKWABAVGCNNNW	183
Qy	1011	AACAGAAATC 1019	
Db	182	DRMAHHHC 174	
RESULT 3			
US-09-621-976-2813/c			
; Sequence 2813, Application US/09621976			
; Patent No. 6639063			
; GENERAL INFORMATION:			
; APPLICANT: Dumas Milne Edwards, J.B.			
; APPLICANT: Jobert, S. J.Y.			
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.			
; FILE REFERENCE: GENSET.054PR2			
; CURRENT APPLICATION NUMBER: US/09/621,976			
; CURRENT FILING DATE: 2000-07-21			
; NUMBER OF SEQ ID NOS: 19335			
; SOFTWARE: Patent.pm			
; SEQ ID NO 2813			
; LENGTH: 832			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: 235..399			
US-09-621-976-2813			
Query Match 3.9%; Score 42.2; DB 4; Length 832;			
Best Local Similarity 11.8%; Pred. No. 0.0086;			
Matches 47; Conservative 172; Mismatches 180; Indels 0; Gaps 0;			
Qy	619	GATATGATAATCTACGTGAGTCTAGTGGCAGTTGTGTAGCGTGGTGGGGCTTTTGGCT	678
Db	408	GATTTGCTAAGTACAGAATTTGACCAATAATATATATTTTGTYYTWWKTWYWT	349
Qy	679	AGCAGTGGGAAAACTTTTGGACAGTGAATCGAATACTCAACAACATCGGAAGGTATCC	738

US-09-657-474-7
; Sequence 7, Application US/09657474
; Patent No. 6399762
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/657,474
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/088,312
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/078,044
; PRIOR FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

Query Match 3.8%; Score 40.6; DB 3; Length 72604;
Best Local Similarity 54.3%; Pred. No. 0.31;
Matches 82; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 288 TCTGACTCCATGGACTCTTACTTACCGTTCTACCTATTCCTGATCTGTGCATC 347
DB 33298 TTGAAATCCATCTGATTCCTCTATTTTCCAGTTGATGCCATGCTCTTGGCCATC 33357

QY 348 TCAGTATAGCCTCAATGCTTCTCTTATTTCTTAACTCACAATACTTACCCCTAT 407
DB 33358 TTTCTTAATTAAGATGCTTCTCTTAACTCTTAACTCTTAACTATTAATCAAGCAGCTATCTACA 33417

QY 408 CATTTTAAATCTCTTTCTCTTAACTATA 438
DB 33418 TATTTTGAATCCCTTAAATCTTGAGCATA 33448

RESULT 7
US-09-806-708B-23/c
; Sequence 23, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1055)
; OTHER INFORMATION: consensus sequence of A.t. and L.a. FAEI promoters
US-09-806-708B-23

Query Match 3.7%; Score 40.4; DB 4; Length 1055;
Best Local Similarity 21.5%; Pred. No. 0.035;
Matches 74; Conservative 99; Mismatches 169; Indels 2; Gaps 1;

QY 150 CTTTCCTGTGCTACTTCCATATATATCTGTCAATTTAAACACATCAACAACTGATAG 209
DB 415 MTTAANNKAGAMWTTAAGNNTWTAATAATKMAAWCARAYCCNNWAAACMAKWKWA 356

QY 210 AGATGGAAGAAACCTCACCTAGGAACCGTGTATTTGGTTTACGTAGTGTGACTTCT 269
DB 355 WTKYAWGAACNNNNKNTAMTCRRRAWYSATWTTWAAATSWWKWYTTTTRKTTMAAANN 296

QY 270 TGTAGGAGCAGATTGCTATCTGTACTCATCTGGACTTCTTACTACCGTTTCTACTA 329
DB 295 NNNNNWAKCKTTSAMWAMWMTWCTCGARTWGGATTMAAACCTAAGAGNSMTYW--CWA 238

QY 330 TTCCCTGATCTGTGCATCTCAGTTAGCTTCAATGCTTTCTCTCTTATTTCTTAACTC 389
DB 237 ATYMGTTWNNNNNNKAWTTTCTMAWCAYTWTARTYKRYTYCYATYTYTSMYMGKM 178

QY 390 AAAAAAATTACCCCTATCATTTTAAATCTCTTTTCTCTTAACTATATCTTCCACCT 449
DB 177 WYWARAAAYASTNCTWSTCRWKTARGWYVAMRYTWAASGNCNNNAKWCAMWYACATN 118

QY 450 ACTTGCAATCAATAATGAGGAGACAGACTCCACAAAAAGTTACAA 493
DB 117 NNAMAWYACAYWMTAGKAAWNTNNKTASGKWMYAMMKTTWYWA 74

RESULT 8
US-09-621-976-18033/c
; Sequence 18033, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano J.V.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18033
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 16
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-18033

Query Match 3.6%; Score 39.2; DB 4; Length 474;
Best Local Similarity 14.3%; Pred. No. 0.052;
Matches 44; Conservative 127; Mismatches 136; Indels 0; Gaps 0;

QY 591 AAAGAAGCAAAATTTCTCAGAAGTTATGGATATGATAATCTACGTGAGTCTAGTGGCCAG 650
DB 307 ARRTWARRAAMYTWITWYMKSGCMWKMWRKYMKCYWWRKYCMRKKRRRG 248

QY 651 TTGTGTTAGCGTGTGGGCTTTTGTCTAGCAGTGTAGGAAAACCTTTGACGACTGAAT 710
DB 247 KKRARSYKSGYKGSYTYKGGKGGTGGKAWSKKRRRAMTTTTYYKKKWRSYWAA 188

QY 711 GGATAACTACAAACATGGGAAGGTATCTCACTATTGAACCTAGTGTGGACAGCTGTAC 770
DB 187 RGKTTKKWAMMMWMMWYMRSMYYKKWAAWKKSCCCWARKKGGSSGRWYWKRG 128

QY 771 CTGCGAGTATTTCTCCATCGGTGGCAGGACTGATCTTCGAGCTCTCTCTATTCTC 830
DB 127 CMRRRRWYSRRRCMWYCTKGSMAWMSGKRAAMCCSKYVYTWCMMAAAWWTWAA 68


```
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
```

```
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664854)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1
```

Query Match 3.5%; Score 37.8; DB 4; Length 1664976;
Best Local Similarity 49.7%; Pred. No. 12;
Matches 96; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

```
Qy 281 ATTGCTATCTGTACTCCATTGGACTTCTTTACTTACCGTTTCTACCTATTCCCTGATCT 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663618 ATTTTTCATCAACCTCATCTGAAATCTGCAAAATTACCAAACTACATCATATA 663677

Qy 341 GTGCATCTCAGTTAGCTTCAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663678 GAGAGGATATTAGCCCTCATCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 663737

Qy 401 CCCCTATCATTTTAAATCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663738 AAACCTGGTTTTTAAATTAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 663797

Qy 461 ATAATGAGGAGAC 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663798 ATATTTAAATAC 663810
```

RESULT 11

```
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
```

```
LOCATION: (84808)..(84808)
FEATURE:
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
```

```
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:40:28 ; Search time 567.648 Seconds
(without alignments)
9996.736 Million cell updates/sec

Title: US-09-913-767-6

Perfect score: 1081
Sequence: 1 caaatccaacagtgtaaat.....ataagctgattttgaaag 1081

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1077.8	99.7	1081	3	Aaa97924 A. thalia
2	1043.8	96.6	1047	6	Adg88190 A. thalia
3	835.8	77.3	1071	3	Aaa97925 A. thalia
4	406.8	37.6	1293	3	Aaa97922 A. thalia
5	406.2	37.6	3387	6	Abz14407 Arabidops
6	406.2	37.6	3387	8	Ada68246 Arabidops
7	369.6	34.2	1194	3	Aaa97923 A. thalia
8	300	27.8	1295	3	Aac37289 Arabidops
9	298.8	27.6	2660	3	Aac47974 Arabidops
10	298.4	27.6	1145	3	Aaa97921 A. thalia
11	298.4	27.6	1292	3	Aac51119 Arabidops
12	234.2	21.7	83698	6	Abn85767 Arabidops
13	225.8	20.9	1403	3	Aac37652 Arabidops
14	222.8	20.6	2175	8	Ada70748 Rice gene
15	211.6	19.6	1152	8	Ada70213 Rice gene
16	145.6	13.5	1049	3	Aaa97920 A. thalia
17	138.6	12.8	592	10	Abx56692 Arabidops
18	136.6	12.6	2586	3	Aac44184 Arabidops
19	136	12.6	1225	3	Aaa97919 A. thalia
20	129.2	12.0	1155	8	Ada69686 Rice gene
21	127	11.7	1040	8	Ada69568 Rice gene

22	126.4	11.7	332	12	ADP92185
23	72.2	6.7	774	6	ABN99127
24	66.6	6.2	477	3	AAC37273
25	66.2	6.1	386	6	ABO85215
26	60.8	5.6	1173	12	ADN72520
27	60	5.6	1200	8	ADA70627
28	57.6	5.3	1421	3	AAC39779
29	56	5.2	1418	3	AAC46395
30	51.6	4.8	2000	8	ADA71938
31	50.4	4.7	493	3	AAC36831
32	47.8	4.4	2000	8	ADA71938
33	47.4	4.4	492	3	AAC36845
34	42	3.9	6067	6	ABN80088
35	42	3.9	110000	6	ABSS5320_0
36	42	3.9	341511	6	ABSS5320
37	40.8	3.8	16633	6	ABN79985
38	40.6	3.8	72604	2	AAZ10752
39	40.6	3.8	72604	6	ABK43231
40	40.4	3.7	4590	5	AAH24065
41	39.8	3.7	264	12	ACH80963
42	39.8	3.7	526	12	ACH67258
43	39.8	3.7	5520	5	AAC68924
44	39.4	3.6	6467	6	ABN80169
45	39.2	3.6	446	3	AAC37188

ALIGNMENTS

RESULT 1
AAA97924
ID AAA97924 standard; DNA; 1081 BP.
AC AAA97924;
XX

DT 19-JAN-2001 (first entry)

XX A. thaliana PUP1 DNA #6.

XX PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.

XX Arabidopsis thaliana.

XX DE19907209-A1.

XX PD 24-AUG-2000.

XX PF 19-FEB-1999; 99DE-01007209.

XX PR 19-FEB-1999; 99DE-01007209.

CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX
SQ Sequence 1081 BP; 322 A; 231 C; 205 G; 323 T; 0 U; 0 Other;

Query Match 99.7%; Score 1077.8; DB 3; Length 1081;
Best Local Similarity 99.8%; Pred. NO. 8e-297;
Matches 1079; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAATCCAAACAGTTCAAGATGAAGAAATTCAGTCAGTAGCAGCAAGCAGAAAGTATCTC 60
DB 1 CAATCCAAACAGTTCAAGATGAAGAAATTCAGTCAGTAGCAGCAAGCAGAAAGTATCTC 60
QY 61 ACTCTAACACATACAAACCGTGGCTCAGGAGTATATCTATGACAAACGAGGAAACAGTAAA 120
DB 61 ACTCTAACACATACAAACCGTGGCTCAGGAGTATATCTATGACAAACGAGGAAACAGTAAA 120
QY 121 TGGCTAGCAACGGTAGTTCAACTTGTGGCTTTCCTGTGCTACTTCCATATATATCTTG 180
DB 121 TGGCTAGCAACGGTAGTTCAACTTGTGGCTTTCCTGTGCTACTTCCATATATATCTTG 180
QY 181 TCATTTAAACACATGCAACCACTGATAGATGGAAGAAAGAACTCACCTAGGAACCGT 240
DB 181 TCATTTAAACACATGCAACCACTGATAGATGGAAGAAAGAACTCACCTAGGAACCGT 240
QY 241 GTATTGGTTTACGTAGTCTTGGACTTCTTGTAGGAGCAGATTGCTATCTGTACTCCATT 300
DB 241 GTATTGGTTTACGTAGTCTTGGACTTCTTGTAGGAGCAGATTGCTATCTGTACTCCATT 300
QY 301 GGACTTCTTTACTTACCCGTTTCTACCTATTCCCTGATCTGTGATCTCAGTTAGCCCTTC 360
DB 301 GGACTTCTTTACTTACCCGTTTCTACCTATTCCCTGATCTGTGATCTCAGTTAGCCCTTC 360
QY 361 AATGCTTTCTCTCTTATTTCTTAACCTCAACAAACCTTACCCCTATCATTTTAAATTTCT 420
DB 361 AATGCTTTCTCTCTTATTTCTTAACCTCAACAAACCTTACCCCTATCATTTTAAATTTCT 420
QY 421 CTTTTCCTCTTAACTATATCTTCCACCTACTTGCATTCATTAATGAGGACAGACTCC 480
DB 421 CTTTTCCTCTTAACTATATCTTCCACCTACTTGCATTCATTAATGAGGACAGACTCC 480
QY 481 ACAAAGTTACAAAGGAGATGTCGAAAGTTTCATATGACCGTTGCTGCTCTGCT 540
DB 481 ACAAAGTTACAAAGGAGATGTCGAAAGTTTCATATGACCGTTGCTGCTGCTCTGCT 540
QY 541 GGTATGCTAGTCTTATCCTCAACACAGCTAGCCCTTCTTAAAGTCTTAAAGAGCAA 600
DB 541 GGTATGCTAGTCTTATCCTCAACACAGCTAGCCCTTCTTAAAGTCTTAAAGAGCAA 600
QY 601 AATTTCTCAGAGTTATCGATATGATATGAACTAGTGTGACAGCTGTTTACCTGGCAGTTA 660
DB 601 AATTTCTCAGAGTTATCGATATGATATGAACTAGTGTGACAGCTGTTTACCTGGCAGTTA 660
QY 661 GTGTGGGGCTTTTGTCTAGCAGTCAGTGGAAACCTTTGACAGTGAATCGATACTAC 720
DB 661 GTGTGGGGCTTTTGTCTAGCAGTCAGTGGAAACCTTTGACAGTGAATCGATACTAC 720
QY 721 AAACATGGGAAGGTATCTCATATATGAACTAGTGTGACAGCTGTTTACCTGGCAGTTA 780
DB 721 AAACATGGGAAGGTATCTCATATATGAACTAGTGTGACAGCTGTTTACCTGGCAGTTA 780
QY 781 TTCTCCATCGGTGGCAGAGGACTGATCTTCGAGCTCTCCTCTTATTTCTCAATATGCAATA 840
DB 781 TTCTCCATCGGTGGCAGAGGACTGATCTTCGAGCTCTCCTCTTATTTCTCAATATGCAATA 840
QY 841 AGCGTTTGGGACTCCAGTGGTTCCTATCTTGGCTGTAATCATTTTCCATGACAAATG 900
DB 841 AGCGTTTGGGACTCCAGTGGTTCCTATCTTGGCTGTAATCATTTTCCATGACAAATG 900
QY 901 AATGGGTTAAAGGTGATTTCTATGATCTTAGCTATTTTGGGGTTTCACATTCCTATGCTCTAC 960
DB 901 AATGGGTTAAAGGTGATTTCTATGATCTTAGCTATTTTGGGGTTTCACATTCCTATGCTCTAC 960

DB 901 AATGGGTTAAAGGTGATTTCTATGATCTTAGCTATTTTGGGGTTTCACATTCCTATGCTCTAC 960
QY 961 CAACAATATCTTGTATGACAAAACTTTGAGAAAAATCATGAAATCACAACAGAAATCC 1020
DB 961 CAACAATATCTTGTATGACAAAACTTTGAGAAAAATCATGAAATCACAACAGAAATCC 1020
QY 1021 CCTGACCCACAGAGCAGAGAGTCAACTTTGGCAATCAAAATAAGCTGATATTTTCAAA 1080
DB 1021 CCTGACCCACAGAGCAGAGAGTCAACTTTGGCAATCAAAATAAGCTGATATTTTCAAA 1080
QY 1081 G 1081
DB 1081 G 1081
RESULT 2
ADG88190
ID ADG88190 standard; cDNA; 1047 BP.
XX AC ADG88190;
XX DT 22-APR-2004 (first entry)
XX DE A. thaliana RPP4-upregulated pathogen infection-related gene #632.
XX KW Pathogen infection-related gene; Plant; Peronospora parasitica;
KW defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;
KW fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX OS Arabidopsis thaliana.
XX PN WO200222675-A2.
XX PD 21-MAR-2002.
XX PF 14-SEP-2001; 2001WO-US028506.
XX PR 15-SEP-2000; 2000US-0322778P.
XX PR 22-JUN-2001; 2001US-0300183P.
XX SY (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PA (UYNC-) UNIV NORTH CAROLINA.
XX PA (GLAZ/) GLAZEBROOK J.
XX PA (WANG/) WANG X.
XX PA (DANGL/) DANGL J L.
XX PA (EULG/) EULGEM T.
XX PA (ZHUT/) ZHU T.
XX GL Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;
XX WPI; 2002-292409/33.
XX Novel isolated polynucleotide, useful for conveying pathogen resistance
XX to plants, and for identifying plants infected with a pathogen.
XX Claim 3; SEQ ID NO 632; 605pp; English.
XX The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC ADG87557) whose expression is altered in response to pathogen infection,
CC and to homologues of these genes from other plants or fungi, especially
CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC expression of genes of the invention was upregulated or downregulated in
CC Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC indicating that they play a role in defence mechanisms. The genes of the
CC invention are regulated by RPP7 or RPP8 which act via unconventional
CC signalling cascades, or by the RPP4-dependent pathway. The invention also
CC relates to polypeptides encoded by the pathogen infection-related genes;
CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC ; expression cassettes, host cells and pathogen-resistant transgenic
CC plants and their progeny comprising a polynucleotide of the invention;
CC and a method of identifying a plant cell infected with a pathogen. The
CC polynucleotide sequences and methods of the invention are useful for

CC identifying plants infected with a pathogen, and for conferring
CC resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC nematodes and insects (e.g., aphids). The present sequence represents an
CC *Arabidopsis thaliana* gene whose expression is altered in response to
CC *Percnospora parasitica* infection. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC [ftp.wipo.int/pub/published pct sequences](http://www.wipo.int/pub/published/pct_sequences).

```
Query Match          96.6%; Score 1043.8; DB 6; Length 1047;
Best Local Similarity 99.8%; Pred. No. 4e-287;
Matches 1045; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy	19	ATGAAAGAAATTCAGTCAGTAGCAGCCAGCAGAGATATCTCATCTTAAACATACAAC	78
Db	1	ATGAAAGAAATTCAGTCAGTAGCAGCCAGCAGAGATATCTCACTCTAAACATACAAC	60
Qy	79	GGTGGCTCAGGAGTATACATATGACAAACGGAGGAAACAGTAATCGCTAGCAACGGTAGTT	138
Db	61	GGTGGCTCAGGAGTATACATATGACAAACGGAGGAAACAGTAATGGCTAGCAACGGTAGTT	120
Qy	139	CAACTGTGTGGCTTTCCCTGTGCTACTTCCCATATTTATATCTTGTCAATTTAAAAACAATGCA	198
Db	121	CAACTGTGTGGCTTTCCCTGTGCTACTTCCCATATTTATATCTTGTCAATTTAAAAACAATGCA	180
Qy	199	ACAACATAGAGATGGAAAGAAACCTCACTAGGAACCGGTATTTGGTTTACGTAGTG	258
Db	181	ACAACATAGAGATGGAAAGAAACCTCACTAGGAACCGGTATTTGGTTTACGTAGTG	240
Qy	259	CTTGGACTCTTTGATAGGAGCAGATTGCTATCTGTACTCCATTTGAGCATCTTTTACTTACCC	318
Db	241	CTTGGACTCTTTGATAGGAGCAGATTGCTATCTGTACTCCATTTGAGCATCTTTTACTTACCC	300
Qy	319	GTTTCTACCTATTCCTCGATCTGTGCATCTCAGTTAGCCTTCAATGCTTCTCTCTTAT	378
Db	301	GTTTCTACCTATTCCTCGATCTGTGCATCTCAGTTAGCCTTCAATGCTTCTCTCTTAT	360
Qy	379	TTTCTTAACTACAAAAACCTTACCCCTATCATTTTTAAATTCCTCTTTTCTCTTAACTATA	438
Db	361	TTTCTTAACTACAAAAACCTTACCCCTATCATTTTTAAATTCCTCTTTTCTCTTAACTATA	420
Qy	439	TCCTCCACCCCTACCTTGCAATTCAAATAGGAGAGACAGACTCCACAAAGTTTACAAAAGGA	498
Db	421	TCCTCCACCCCTACCTTGCAATTCAAATAGGAGAGACAGACTCCACAAAGTTTACAAAAGGA	480
Qy	499	GAGTATGCTCAAAGTTTTCATATGCACCGTTGCTGCGTCTGCTGTTATGGTCTAGTCTTA	558
Db	481	GAGTATGCTCAAAGTTTTCATATGCACCGTTGCTGCGTCTGCTGTTATGGTCTAGTCTTA	540
Qy	559	TCCTCAACAGCTAGCGCTTTCTAAAAGTCTTAAAGAGCAAAATTTCTCAGAAAGTTATG	618
Db	541	TCCTCAACAGCTAGCGCTTTCTAAAAGTCTTAAAGAGCAAAATTTCTCAGAAAGTTATG	600
Qy	619	GATATGATAATCTACGTGAGTCTAGTGGCCAGTTGTTTACGTGTTGGGGCTTTTGGCT	678
Db	601	GATATGATAATCTACGTGAGTCTAGTGGCCAGTTGTTTACGTGTTGGGGCTTTTGGCT	660
Qy	679	AGCAGTCAGTGGAAAACTTTTGACAGTGAATGGATAACTACAAACATGGGAAGGTATCC	738
Db	661	AGCAGTCAGTGGAAAACTTTTGACAGTGAATGGATAACTACAAACATGGGAAGGTATCC	720
Qy	739	TACATTTAGAACTAGTGTGGACAGCTGTTCCTGGCAGTTATTCTCCATCGGTGGGACA	798
Db	721	TACATTTAGAACTAGTGTGGACAGCTGTTCCTGGCAGTTATTCTCCATCGGTGGGACA	780
Qy	799	GGACTGATCTTCGAGCTCTCTCTCTATCTTCAAAATGCAATPAACGGTTTTGGGACTCCCCA	858
Db	781	GGACTGATCTTCGAGCTCTCTCTCTATCTTCAAAATGCAATPAACGGTTTTGGGACTCCCCA	840
Qy	859	GTGGTTCCTATCTTGGCTGTAATCATTTTCCATGACAAAATGAAATGGGTATAAGGTGATT	918

Db	841	GTGGTTCTATCTTGGCTGTAATCATTTTCCATGACAAAATGAATGGGTTAAAGGTGATT	900
Qy	919	TCATATGATCCTAGCTATTTCGGGGTTTCACATTCCTATGCTACCAACAATAATCTTTGATGAC	978
Db	901	TCATATGATCCTAGCTATTTCGGGGTTTCACATTCCTATGCTACCAACAATAATCTTTGATGAC	960
Qy	979	AAAAACTTGAGAAAAAATCATGAAATCACAACAACAGAATCCCTCTGACCCACCAAGAAGCA	1038
Db	961	AAAAACTTGAGAAAAAATCATGAAATCACAACAACAGAATCCCTCTGACCCACCAAGAAGCA	1020
Qy	1039	GAAGAGTCAACTTGGCAATCAAAAATTA	1065
Db	1021	GAAGAGTCAACTTGGCAATCAAAAATTA	1047

RESULT 3
AAA97925
ID AAA97925 standard; DNA; 1071 BP.
XX
XX AC AAA97925;
XX
XX 19-JAN-2001 (first entry)
XX
XX A. thaliana PUP1 DNA #7.
XX
XX PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.
XX
XX Arabidopsis thaliana.
OS
XX DE19907209-A1.
PN
XX
XX 24-AUG-2000.
PD
XX
XX 19-FEB-1999; 99DE-01007209.
PF
XX
XX 19-FEB-1999; 99DE-01007209.
PR
XX
XX (FROM/) FROMMER W.
PA
XX
XX Gillissen B, Buerkle L, Andre B, Frommer WB;
PI
XX WPI; 2000-566202/53.
DR
XX
XX Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transport, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.
PT
XX
XX Claim 2; Page 15; 24pp; German.

XX This invention describes a novel nucleic acid encoding a plant nucleobase
CC transporter (I). (I) is produced by complementation of a nucleobase
CC transport (NBT)-defective host cell with a plant gene bank by selection
CC of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
CC expression of (II) (when in antisense orientation), and to produce
CC transgenic crop plants. The transgenic plants have modified nucleobase
CC transport properties, e.g. altered affinity and substrate specificity
CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX
XX Sequence 1071 BP; 307 A; 236 C; 208 G; 320 T; 0 U; 0 Other;
SQ

	Query Match	77.3%	Score 835.8	DB 3	Length 1071
	Best Local Similarity	90.2%	Pred. No. 8.9e+228		
	Matches 894	Conservative 0	Mismatches 97	Indels 0	Gaps 0
QY	89	GAGTATCTATGACACACGGAGGAAAAACAGTAAATGGCTAGCAACGGTAGTTTCAACTTGTTC	148		
Db	80	GACTTACTATGAAATGAGGAGAAACAGCAATATGGCTAGCAACGGTAGTTTCAACTTGTAG	139		

Db 653 CTGCGCTCTCGTGGTCAACACTGATTCGGAAAAACACAGCAAAAGTGTCTAGAGTAAAT 712
Qy 503 ATGTCAAGGTTTTCATATGACCGTGTCTGCGTCTGCTGTTATGTGTAGTCTTATCCC 562
Db 713 ATGTGATAGGATAATATGATACCATTTGGTGTCTCTGCTGGATTGATGCTGCTATCCC 772
Qy 563 TACAACAGCTAGCTCTTCTAAAGTCTCTAAAGAGCAAAATTTCTCAGAAAGTTATGGATA 622
Db 773 TGTGTACAACTGATCTCTCAGGAAGGTTTAAAGAGCAAAATTTCTCAACGGTCACTGACT 832
Qy 623 TGATAATCTACGTAGCTAGTGGCCAGTGTGTAGCTGTGGGCTTTTGTCTAGCA 682
Db 833 TGTGTGCTTACCAATCTCTAGTTGCAAGCTGTGTGTCTCATAGGACTTTTTCGCAAGCG 892
Qy 683 GTGAGTGGAAAACTTTGAGCAGTGAATGATTAATCAAAACATGGGAAGGTATCCTACA 742
Db 893 GGGAGTGGAAAACTTTAAACAAGTGAGTGGAAAACTTACAACTGGGGAAGTGCATAGC 952
Qy 743 TTATGAACCTAGTGTGGACAGCTGTTTACCTGGCAGTTATTTCTCCATCGGTGGCAGAGAC 802
Db 953 TTATGACTTTGGGCTCGATAGCTATTTCTCTGCAAGTGTGTGTCTTCAAGCTTACACCA 1012
Qy 803 TGATCTTCGAGCTCTCTCTCTATTTCTCAATGCAATAGCGTTTGGGACTCCGAGTGG 862
Db 1013 TGATCTTTGAGTCATCTTCTGTGTCTTCTCAATTTCCAAATCTGCTGTGGATTGCTTATAG 1072
Qy 863 TTCTATCTTGGCTGTAAATCATTTTCCATGACAAAATGAATGGTTTAAAGGTGATTTCTA 922
Db 1073 TTCCAGTTGTAGCAGTGTGTTTTCATGATTAATGAACGGTCAAAAGATCTTCTCCA 1132
Qy 923 TGATCTAGCTATTTGGGTTTCACTTCTCTATGTTTACCAACAATATTTGATGACAAA 982
Db 1133 TCATTTTAGCTATCTGGGATTCTATTTCTGTTCTATCAGCACTACCTCGAGAAAGA 1192
Qy 983 ACTTGAAGAAAATCATGAATCAACAACAGATCCCTGATCCACCAAGAGCAAGAG 1042
Db 1193 AGTTGAAGACTAGCCACAAAGTCTGTAGGAGATCTCATCTACTACTGCTGTGAGGAAG 1252

RESULT 5

ABZ14407
ID ABZ14407 standard; DNA; 3387 BP.

XX
AC ABZ14407;

XX
DT 21-JAN-2003 (first entry)

XX
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 2212.

XX
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX
OS Arabidopsis thaliana.

XX
PN WO200216655-A2.

XX
PD 28-FEB-2002.

XX
PF 24-AUG-2001; 2001WO-US026685.

XX
PR 24-AUG-2000; 2000US-0227866P.

XX
PR 26-JAN-2001; 2001US-0264647P.

XX
PR 22-JUN-2001; 2001US-0300111P.

XX
XX (SCRI) SCRIPPS RES INST.

XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX
PI Harper JF, Kreps J, Wang X, Zhu T;

XX
XX WPI; 2002-304127/34.

XX
XX Identifying a stress condition to which a plant cell has been exposed and

PT
PT producing plants with increased tolerance to these abiotic stresses.

XX

PS Claim 144; SEQ ID NO 2212; 577pp + Sequence Listing; English.

XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office

XX
SQ Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;

Query Match 37.6%; Score 406.2; DB 6; Length 3387;

Best Local Similarity 64.7%; Pred. No. 5.5e-105;

Matches 622; Conservative 0; Mismatches 333; Indels 6; Gaps 1;

Qy 88 GGAGTATATGACAAACGAGGAAACAGTAATGGCTAGCAACGGTAGTTCACATTGTT 147

Db 226 GTAGTTTACTATGAAATGGTGGAAAGATACATGGATGGGAACACTTGTCCAACCTAATC 285

Qy 148 GGCCTTTCCTGCTACTTCCATATATATATCTTGTCAATTTAAACACATGCAACACTGAT 207

Db 286 GGCCTTCCCTGTTCTGTTCTCTTCGCTTCTTTTCCCAACCAAAATCCCAACCAACA 345

Qy 208 AGAGATGGAAAAAGAACCTCACCT-----AGGAACCGTGTATTTGGTTTACGTAGTCTT 261

Db 346 GAAGCAGATTTTCAGAAAGTTCTCTTCTTCCATTCTTGGATCAGTTTACATCGTTACT 405

Qy 262 GGACTTCTTTAGGAGCAGATGCTATCTGTACTCCATTTGGACTTCTTTTACTTACCCGTT 321

Db 406 GGACTATTAGTGTCTGCTAACTCTTATATGCTCTGTGTGTTTACTATATCTTACAGTT 465

Qy 322 TCTACTATTCCCTCATCTGTGCATCTCAGTTTGTAGCTTCAATGCTTCTTCTCTTATTTT 381

Db 466 TCTACTTCTTCCCTCATCTGGCCTCAATTTGGCCTTCACTGCTTCTTCTCATATTTT 525

Qy 382 CTTAACACAAAACTTACCCCTATCATTTTAAATTTCTTTTCTCTTAACTATATCT 441

Db 526 CTAAACTCGAGAAGTTCACACCTTTTCATTTGGAATTTCTGTCTTCTTCTTATTTCC 585

Qy 442 TCCACCTTACTTGCATTTCAATATGAGGAGACAGACTCCACAAAAAGTTTACAAAAGAGAG 501

Db 586 TCTGCCCTCTCGTGTCAACACTGATTCGGAAAAACACAGCAAAAGTGTCTAGAGTAAAA 645

Qy 502 TATGTCAAAGGTTTCATATGACCGTTCGCTGCTGCTGTTATGGTCTAGTCTTATCC 561

Db 646 TATGTGATAGGATTAATATGATACCATTTGGTGTCTCTGCTGGATTGGATTGCTGCTATCC 705

Qy 562 CTACAACAGCTAGCCTTTTCTAAAGTCTCTAAAGAGCAAAATTTCTCAGAAAGTTATGGAT 621

Db 706 CTGGTACAACATGATCTCTCAGGAAGTTTAAAGAAGCAAACTTCTCAACGCTCACTGAC 765

Qy 622 ATGATAATCTACGTGAGTCTAGTGCCAGTGTGTGTAGCGGTGGGGCTTTTTCGCTAGC 681

Db 766 TTGGTTCGCTTACCAATCTCTAGTTGCAAGCTGTGTGTTTCTCATAGGACTTTTTCGCAAGC 825

Qy 682 AGTGAGTGGAAAACTTTTCAGCAGTGAAATGATTAATACTCAACATGGGAAGTATCCTTAC 741

Db 826 GGGGAGTGGAAAACTTTTAAAGAGTGAATGAAAACTCAAACTGGGGAAGTGCATATAC 885

Qy 742 ATTATGAACCTAGTGTGACAGCTGTTTACCTGGCAGTTTATTTCTCCATCGGTGGCAGAG 801

Db 886 GTTATGACTTTGGCCTCGATAGCTATTTCTGGCAAGTCTACACATTTGGGCTCGTGGGA 945

Qy 802 CTGATCTTCGAGCTCTCTCTCTATTTCTCAATGCAATGAAAGGTTTTCGGAATCCCAAGT 861

Db 946 CTGATCTTTGAGTCACTCTCTCTGTTCTTCCAAATTCCTCAATGCTGTGGGATTGCTATA 1005

DT	19-JAN-2001	(first entry)
XX		
DE	A. thaliana PUP1 DNA #5.	
XX		
KW	PUP1; transgenic plant; nucleobase transporter; apical dominance;	
KW	flowering behaviour; senescence; pesticide distribution; ds.	
XX		
OS	Arabidopsis thaliana.	
XX		
FN	DE19907209-A1.	
XX		
PD	24-AUG-2000.	
XX		
PF	19-FEB-1999; 99DE-01007209.	
XX		
PR	19-FEB-1999; 99DE-01007209.	
XX		
PA	(FROM/) FROMMER W.	
XX		
FI	Gillissen B, Buerkle L, Andre B, Frommer WB;	
XX		
DR	WPI; 2000-566202/53.	
XX		
PT	Nucleic acid, useful for producing transgenic plants with altered	
PT	nucleobase transporter, encodes a nucleobase transporter protein of	
PT	Arabidopsis thaliana.	
XX		
PS	Claim 1f; Page 14; 24pp; German.	
XX		
CC	This invention describes a novel nucleic acid encoding a plant nucleobase	
CC	transporter (I). (I) is produced by complementation of a nucleobase	
CC	transport (NBT)-defective host cell with a plant gene bank by selection	
CC	of NBT-positive cells. (I) is used to isolate homologous sequences from	
CC	bacteria, fungi, plants, animals and humans, for expression of the	
CC	encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting	
CC	expression of (II) (when in antisense orientation), and to produce	
CC	transgenic crop plants. The transgenic plants have modified nucleobase	
CC	transport properties, e.g. altered affinity and substrate specificity	
CC	that may result in more efficient nucleobase transport in leaves, changes	
CC	in apical dominance, flowering behaviour and senescence, or improved	
CC	distribution of pesticides. This sequence encodes the Arabidopsis	
CC	thaliana PUP1 protein which is described in the method of the invention	
XX		
SQ	Sequence 1194 BP; 298 A; 279 C; 231 G; 386 T; 0 U; 0 Other;	
Query Match 34.2%; Score 369.6; DB 3; Length 1194;		
Best Local Similarity 62.2%; Pred. No. 9.9e-95;		
Matches 60; Conservative 0; Mismatches 359; Indels 6; Gaps 1		
QY	89 GAGTATACTATGACAAACGGAGGAACAGTAATGGCTAGCAACGGTAGTTCAACTTGTG 148	
Db	197 GATTGTACTATGAAATGGAGGAAATAGCACATATGTGTGAACACTTCTTCACTCAT 256	
QY	149 GCTTTCCTGTGCTACTTCCATATTATATCTTGTCAATTA-----AAACACATGCAACAA 202	
Db	257 GCTTTCCTGTGCTACTTCCATATTATATCTTGTCAATTA-----AAACACATGCAACAA 316	
QY	203 CTGATAGAGATGGAAAAAGAACCTCACCTAGGAACCGGTATTGGTTTACCTAGTGCCTG 262	
Db	317 ATACAAATTCAGTCAGTCGCCCTTCCCTTTCACACCCCTTGCATCGGTTTACTTGTGCAC 376	
QY	263 GACTTCTGTAGGAGCAGATGCTATCTGTACTCCATTTGGACTTCTTTTACTTACCCGTTT 322	
Db	377 GACTGTAGTGTCCGCTATGCTATTGTCTCGAGTAGGGTTGCTCTACTTACAGTCT 436	
QY	323 CTACCTATTCCTGATCTGTGCGATCTCAGTTAGCCTTCAATGCTTTCTTCTTATTTTC 382	
Db	437 CTACTTTCCTCATCTTGGCCCTCACAGTTGGGCTTCACATGCCTTTTCTCATATTTCC 496	
QY	383 TTAACCTACAAAACTTACCCCTATCATTTTAAATTCCTTTTCTCTTCTTACTATATCTT 442	
Db	497 TTAACCTACAAAAAGTTCACTCCCTTGTAGTACGTTCTTGTCTCTCTCTGATATCTT 556	

PR	01-APR-1999;	99US-0137462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-01351124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135639P.
PR	25-MAY-1999;	99US-0136081P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	06-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	25-AUG-1999;	99US-0149930P.
PR	26-AUG-1999;	99US-0150566P.
PR	27-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.

PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159337P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	26-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 27.8%; Score 300; DB 3; Length 1295;		
Best Local Similarity 58.4%; Pred. No. 7.4e-75;		
Matches 572; Conservative 0; Mismatches 390; Indels 18; Gaps 2;		
QY	89	GAGTATACATGACACAGGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCAACTTGTG 148
DB	192	GACTTTACTAGAAAGCGGTAAAGCNCATGGCTCGAACCTTGGTTCAGCTTGTAG 251
QY	149	GCTTTCCTGTGCTACTTCCATATATATCTGTGCTATTTAAACACATGCAACAACTGATA 208
DB	252	GGTTTCCTTTAAACCTTCTCTGTATATTTACTTAAAGCGCTGACCGTCCAGACTATAAA 311
QY	209	GAGATGGAAGAAAGAACTCACTAG-----GAAACGTGTATGGTTTACGTAGTCTTG 262
DB	312	CCATTACCAAAAATACTTCTCTCTGACACTATCTTTAGTGTATATTTGGACTTG 371
QY	263	GACTTCTGTGAGGACAGATTGCTATCTGTACTCCATTTGACTTCTTTACTTACCGTTT 322
DB	372	GCTTGTCTGTGTGACATGTATTTGTACTCATTTGGGCTACTTTACCTTCTGTCT 431
QY	323	CTACCTATTCCCTGATCTGTGCATCTCAGTTAGCCTTCAATGCTTTCTTCTCTATTTC 382
DB	432	CAACTTCTCTTTGATCTCTGGCTCGCAATTGGCTTTTAAAGCGCTTCTCTTACTTCC 491
QY	383	TAACTCAGAAAACTTACCCTATCATTTTAAATCTCTTTTCTTCTTAACTATATCTT 442
DB	492	TAACTCAGAAAAATCACACACTTTATATCTCAATTCACCTTGTCTCTTAAACATATCTT 551
QY	443	CCACCTCTCTGCTTCAATTAATGAGGACAGACTCCCAAAAGTTACAAAGGAG--- 499
DB	552	CTACCTTCTGTGTTTCCACATGAACCGAATCTCCCTCTTCTACTTCAAAAGTCGCGAG 611
QY	500	-----AGTATCAAGAGTTTTCATATGCAACCGTTGCTGCTGCTGTTATGTC 550
DB	612	CCAAGTCCAAGTATGTGATGGATACATCTGCGCGCTCGGTAGCTCAGCTGTATTCTC 671
QY	551	TAGTCTTATCCCTACAAAGCTAGCTTTCTTAAAGTCTTAAAGTCTTAAAGAAATTTCTCAG 610
DB	672	TGCTGCTTTCTTTAAACAGATTAGCGCTTCGAAAGATTCTTAAAGAAATACATCTCAAG 731
QY	611	AGTTATGATATGATAATCTACGTGAGTCTAGTGGCCAGTTGTGTAGCGTGGGGC 670
DB	732	CTATTTTAGACATGGCCACATATCCGCTCTATGGTAGCTACTTGTGTAGTTGGTAGGAC 791

QY	671	TTTTTGCTAGCAGTGAAGAAACCTTTGAGCAGTGAATGGATAACTACAACATGGGA 730
DB	792	TTTTTGGAAAGTGGTGGTGGAAAAAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGGA 851
QY	731	AGGTATCCTACATATTGAACCTAGTGTGACAGCTGTTACCTGGGAGTTATTCTCCATCG 790
DB	852	AAAGCTCATACATTTTGATAAACATCGGTTCAACGATATCATGGCAAGCTTGTGTTGATTG 911
QY	791	GTGGCACAGGACTGATCTTCGAGCTCTCTCTCTATTTCTCAAAATGCAATGAAGCGTTTGG 850
DB	912	GAAGTGTGTTGATTATCGAAGTTTCATCGCTTTTCCCAATGTCTATAGCACTCTTT 971
QY	851	GACTCCAGTGGTTCCTATCTTTGGCTGTAATCATTTTCCATGACAAAAATGAATGGGTTAA 910
DB	972	GTTTACCAGTGTGCTGTTCTTGTCTGTTGTTCTTCTCGTGATGAGATGAGTGGGAATCA 1031
QY	911	AGGTGATTTCTATGATCCTAGCTATTTGGGGTTTCACTTCTCTATGTCTACCAATATC 970
DB	1032	AGTTGATTGCAATGTTTTTGGCCATCTGGGGATTTGTTTCTTATGGTTATCAGCATATG 1091
QY	971	TTGATGACAAAACTTGAAGAAAAATCATGAAATCACAACACAGAAATCCCTGACCCAC 1030
DB	1092	TCAATGATAGAAAGCCAGAGAGACCAAGAGCTTCTCAGTCTAAAGAGAGAAGAAC 1151
QY	1031	CAGAAGCAGAGAGTCAACT 1050
DB	1152	AAAAACAAGTAGATACCAT 1171
RESULT 9		
AAC47974		
ID	AAC47974 standard; DNA; 2660 BP.	
XX	AC AAC47974;	
XX	DT 18-OCT-2000 (first entry)	
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 55799.	
DE	Hybridisation assay; genetic mapping; gene expression control;	
KW	protein identification; signal transduction pathway; metabolic pathway;	
KW	promoter; termination sequence; ss.	
XX	Arabidopsis thaliana.	
XX	EP1033405-A2.	
XX	PD 06-SEP-2000.	
XX	25-FEB-2000; 2000BP-00301439.	
PF	25-FEB-1999; 99US-0121825P.	
PR	05-MAR-1999; 99US-0123180P.	
PR	09-MAR-1999; 99US-0123548P.	
PR	23-MAR-1999; 99US-0125788P.	
PR	25-MAR-1999; 99US-0126264P.	
PR	29-MAR-1999; 99US-0126785P.	
PR	01-APR-1999; 99US-0127462P.	
PR	06-APR-1999; 99US-0128234P.	
PR	08-APR-1999; 99US-0128714P.	
PR	16-APR-1999; 99US-0129845P.	
PR	19-APR-1999; 99US-0130077P.	
PR	21-APR-1999; 99US-0130449P.	
PR	23-APR-1999; 99US-0130510P.	
PR	23-APR-1999; 99US-0130891P.	
PR	28-APR-1999; 99US-0131449P.	
PR	30-APR-1999; 99US-0132048P.	
PR	30-APR-1999; 99US-0132407P.	
PR	04-MAY-1999; 99US-0132484P.	
PR	05-MAY-1999; 99US-0132485P.	
PR	06-MAY-1999; 99US-0132486P.	
PR	06-MAY-1999; 99US-0132487P.	
PR	07-MAY-1999; 99US-0132863P.	

PR 11-MAY-1999;	99US-0134256P.	PR 23-JUL-1999;	99US-0145218P.
PR 14-MAY-1999;	99US-0134218P.	PR 23-JUL-1999;	99US-0145224P.
PR 14-MAY-1999;	99US-0134219P.	PR 26-JUL-1999;	99US-0145276P.
PR 14-MAY-1999;	99US-0134221P.	PR 27-JUL-1999;	99US-0145913P.
PR 14-MAY-1999;	99US-0134370P.	PR 27-JUL-1999;	99US-0145918P.
PR 18-MAY-1999;	99US-0134768P.	PR 27-JUL-1999;	99US-0145919P.
PR 19-MAY-1999;	99US-0134941P.	PR 28-JUL-1999;	99US-0145951P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146386P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146388P.
PR 24-MAY-1999;	99US-0135629P.	PR 02-AUG-1999;	99US-0146389P.
PR 25-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147038P.
PR 27-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147204P.
PR 28-MAY-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147192P.
PR 03-JUN-1999;	99US-0137528P.	PR 05-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 31-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.

```
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 27.6%; Score 298.8; DB 3; Length 2660;
Best Local Similarity 59.0%; Pred. No. 2.2e-74;
Matches 532; Conservative 0; Mismatches 367; Indels 3; Gaps 1;

QY 92 TATACATGACACGAGGAGAAACAGTAATGGCTAGCAACGGTAGTTCAACTTGTGGCT 151
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
264 TTACTATGATGAAGGTGGAAACAGTAATGGAGCAACTCTGTTCAAACGGCTGCTT 323
QY 152 TTCTGTGCTACCTCCATATATATCTGTGTCATTTAAACACATGCAACAACTGATAGAG 211
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
324 TTCTATCTCTATATCCCGCTT---TTGCTACTTCGGTCTCGGCTAGTAGTCTT 380
QY 212 ATGGAAAAAGAACTCACCTAGGAACCGTGTATTTGGTTTACGTAAGTCTTCTTGG 271
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
381 CGGAGAGTTCGTCTTCACTCAAGTACATTTGTTGATCTATGTTTGTCTCGGTGATCA 440
QY 272 TAGGACGAGATTGCTATCTGTAATCCATTTGGACTTCTTACCTACCGTTTCTACCTAT 331
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
441 TTGCTGGAGATAATATGTTATATCTGTTGGACTTTTGTACCTCTCTGCATCGACGTAT 500
QY 332 CCCTGATCTGTGATCTCAGTTAGCTTCAATGCTTTTCTTCTTAACTATATCTCCACCTAC 391
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
501 CGCTCATTTGCGTACTCAGTTAGCTTCAACCGGTGTTCTTATTTCACTCAATGCTC 560
QY 392 AAAAATTTACCCCTATCATTTAAATTTCTTTTCTTCTTAACTATATCTCCACCTAC 451
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
561 AGAAGTTCACTGTTTGAATTTCACTCCGTTGTTCTTCTTGTCTCGCTGCTTGA 620
QY 452 TTGCATTCATATAGGAGACAGACTCCACAAAATTTACAAAGGAGATGTCMAAG 511
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
621 TAGCTCTCAATGATGATCGGATPACTCTTCTGTGTCTCCAGTCTAAGTATATTTG 680
QY 512 GTTTCATATGACCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 571
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
681 GGTGTTGTGTACACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 740
QY 572 TAGCCTTTCTAAAGTCTTAAAGAACAAATTTCTCAGAAGTTATGATATGATAATCT 631
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
741 TCTCGTTCGAGAAGATTTCTGAAGAGGAGACGTTTCTGTGTTCTTGAATGCAATCT 800
QY 632 ACGTAGTCTAGTGGCAGTTGTTAGGCTGTGTTAGGCTGTGTTAGGCTGTGTTAGGCTGTG 691
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
801 ACATCTCTTTAGTGGCGACTTGTGTTTTCGTTTATAGGCTTTTGTAGCGGGAATGA 860
QY 692 AAATTTTGACGAGTGAATGATTAACATCAACATGGGAAGTATCTCATATATGAACC 751
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
861 GAACGCTGATGGGAATGGAAGGTTATCATAAAGGGAAGCCCTTATGTAAGTACCT 920
QY 752 TAGTGTGGACAGCTGTTACCTGGCAGTTATTTCTCCATCCGTTGGCAGAGCTGATCTCG 811
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
921 TGTCTGGACAGAGTACTTGGCAGTGTGTTCTGTTGAGTGTGTTGAGTGTGTTGATATTC 980
QY 812 AGCTCTCTCTTATTTCTCAATGCAATGAGCGTTTGGGATCCCGAGTGTTCCTATCT 871
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
981 TGTGACGTCGCTCTTCTCAACGTCATTTAGTACGCTCTCTCTAGCTGTGTAATCCACTAG 1040
QY 872 TGGCTGTAATCATTTTCCATGACAAATGAATGGGTTAAAGGTGATTTCTATGATCCCTAG 931
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1041 CAGCTTTGTTGTTTCCGTGATAAATGAGTGGTGTAAAGATTATGGCAATGCTGATCG 1100
```

```
QY 932 CTATTGGGGTTTCACTTCTCTATGCTACCAACAATATCTTGATGACAAAAAATTGAAGA 991
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1101 CTATTGGGGTTTTCGCTTCTTATGTTTACCAGAAATCATATTTGATGACTTGAAGTAAGAC 1160
QY 992 AA 993
DB |||||
1161 AA 1162

RESULT 10
AAA97921
ID AAA97921 standard; DNA; 1145 BP.
XX
AC AAA97921;
XX
DT 19-JAN-2001 (first entry)
XX
DE A. thaliana PUP1 DNA #3.
XX
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.
XX
OS Arabidopsis thaliana.
XX
PN DE19907209-A1.
XX
PD 24-AUG-2000.
XX
PF 19-FEB-1999; 99DE-01007209.
XX
PR 19-FEB-1999; 99DE-01007209.
XX
PA (FROM/) FROMMER W.
XX
PI Gillissen B, Buerkle L, Andre B, Frommer WB;
XX
WP1; 2000-566202/53.
XX
Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transport, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.
XX
Claim 1f; Page 12-13; 24pp; German.
XX
This invention describes a novel nucleic acid encoding a plant nucleobase
transporter (I). (I) is produced by complementation of a nucleobase
transport (NBT)-defective host cell with a plant gene bank by selection
of NBT-positive cells. (I) is used to isolate homologous sequences from
bacteria, fungi, plants, animals and humans, for expression of the
encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
expression of (II) (when in antisense orientation), and to produce
transgenic crop plants. The transgenic plants have modified nucleobase
transport properties, e.g. altered affinity and substrate specificity
that may result in more efficient nucleobase transport in leaves, changes
in apical dominance, flowering behaviour and senescence, or improved
distribution of pesticides. This sequence encodes the Arabidopsis
thaliana PUP1 protein which is described in the method of the invention
XX
Sequence 1145 BP; 309 A; 234 C; 226 G; 376 T; 0 U; 0 Other;
SQ

Query Match 27.6%; Score 298.4; DB 3; Length 1145;
Best Local Similarity 58.3%; Pred. No. 2e-74;
Matches 571; Conservative 0; Mismatches 391; Indels 18; Gaps 2;

QY 89 GAGTATACCTATGACAAACGAGGAAACAGTAATGGCTAGCAACGGTAGTTCAACTTGTG 148
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
140 GACTTTACTAGCAAAAAGCGGTAAAGCACATGCTCGAAACCTTGGTTGAGTTGAGT 199
QY 149 GCTTTCCTGCTACTTCCATATATATCTTGTCTATTTAAACACATCAACAACTGATA 208
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
200 GGTTCCTTTTAAACCTTCCCTTGTCTATTTACTTAAAGCCTGAGCCGCTCAAGACTAAA 259
QY 209 GAGATGGAAAAAGAACCTCACCTAG-----GAACCGTGTATTGTTTACGTAAGTGTG 262
```

Db 260 CAAATTACCAAAAACTACTTCTCCCTCTTGACACTATCTTTAGTGATATATGGACTTG 319
Qy 263 GACTTCTTGTAGGACAGATGCTATCTGTACTCCTCAATGGACTTCTTTTACATCCCGTTT 322
Db 320 GCTTCTTGTGTGCTGGACAGATGATATTTGTACTCATTTGGGCTACTTTACCTTCTGTCT 379
Qy 323 CTACCTATTCCCTGATCTGTGCTGATCTCAATGCTTCAATGCTTCTTCTTCTTATTTTC 382
Db 380 CAACCTTCTCTTTGATCTCTGGTCGCAATTTGGCTTTTAAACCGCTCTCTTACTTTC 439
Qy 383 TTAACCTACAAAACTTACCCCTATCATTTTAAATTTCTCTTTCTCTTAACTATATCTT 442
Db 440 TAAACTCACAAAAATCACACCAATTTATCACTTCACTTGTCTCTTACCATACTT 499
Qy 443 CCACCTACTTGCATTCATTAATATAGGAGACAGACTCCACAAAAATTTACAAAAGAG--- 499
Db 500 CTACACTTCTTGTATCAACATGAACCAAGATCTCCCTCTTCTACTTCAAAAGTCCGCAG 559
Qy 500 -----AGTAGTCAAGGTTTCATATGCACCGTTGCTCGCTCTGCTGTTATGTC 550
Db 560 CCAAGTCCAAGTATGTGATGATACATCTGCGCGTGGTAGCTCAGCTGGTTATTCTC 619
Qy 551 TAGTCTTATCCCTACAAACAGCTAGCTTTCTTAAAGTCTTAAAGACCAAAATTTCTCAG 610
Db 620 TGGTCTTCTTTAACAGATTACGGTTTCGAAAGATTTCTAAGAAATACACATTCAGG 679
Qy 611 AAGTTATGATATGATATCTACGTGAGTCTAGTGGCCAGTTGTGTAGCOTGTGGGGC 670
Db 680 CTAATTTAGACATGGCCACATATCGTCTATGTTAGTCTACTTGTGTAGTTGTGGTAGGAC 739
Qy 671 TTTTGTCTAGCAGTGAAGTGAACCTTTGACGAGTGAATGATTAATACAAACATGGGA 730
Db 740 TTTTGGAAAGTGGTGGGTGGAAGAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGGGA 799
Qy 731 AGGTATCCTACATTAATGACCTAGTGTGGACAGCTGTTTACCTGGCAGTTATTTCCATCG 790
Db 800 AAAGCTCATACATTTTGTATAACATCGTTCAACGATATCATGGCAAGCTGTTTGATTG 859
Qy 791 GTGGCACAGGACTGATCTTCGAGCTCTCTCTTATCTTCAATTCAAATGCAATAGCGTTTGG 850
Db 860 GAAAGTGTGGTTGATATCGAAGTTTCATCGCTTTTTCCTCAATGTCATAAGCACTCTT 919
Qy 851 GACTCCCAAGTGGTCTTCTTCTGCTGTAATCATTTTCCATGACAAATGAATGGTTAA 910
Db 920 GTTTACCAAGTGTGCTGCTTCTTCTGTTGTTCTTCCGTGATGAGATGGAATCA 979
Qy 911 AGGTGATTTCTATGATCCTAGCTATTTGGGGTTTCACTTCCCTATGCTTACCAACAATATC 970
Db 980 AGTTGGTTGCAATGTTTTTGGCCATCTGGGATTTGTTTCTTATGGTTATCAGCATATG 1039
Qy 971 TTGATGACAAAACTTGAAGAAAAATCATGAAATCAACAAACAGAAATCCCTGACCCAC 1030
Db 1040 TCAATGATGAAGAGCCAGAAGAACCAAGAGCTTCTCAGTCTTAAAGAAGAAGAAC 1099
Qy 1031 CAGAGCAGAGAGTCAACT 1050
Db 1100 AAAAAACAAGTAGATACCAATT 1119

RESULT 11
ID AAC51119 standard; DNA; 1292 BP.
XX
AC AAC51119;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 67344.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.

XX OS Arabidopsis thaliana.
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PF 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.

PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.	Query Match	27.6%;	Score 298.4;	DB 3;	Length 1292;
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.	Best Local Similarity	58.3%;	Pred. No. 2.1e-74;		
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.	Matches 571;	Conservative	0;	Mismatches 391;	Indels 18;
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.					Gaps 2;
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.					
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.					
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.					
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.					
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.					
PR	13-JUL-1999;	99US-0143342P.	PR	04-OCT-1999;	99US-0157117P.					
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.					
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.					
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.					
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.					
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.					
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.					
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159294P.					
PR	19-JUL-1999;	99US-0144334P.	PR	13-OCT-1999;	99US-0159295P.					
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159329P.					
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159330P.					
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159331P.					
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.					
PR	21-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.					
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.					
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.					
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.					
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.					
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.					
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.					
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.					
PR	23-JUL-1999;	99US-0145445P.	PR	22-OCT-1999;	99US-0160980P.					
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.					
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.					
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.					
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.					
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.					
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.					
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.					
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.					
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.					
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.					
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.					
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.					
PR	04-AUG-1999;	99US-0147302P.								
PR	05-AUG-1999;	99US-0147192P.								
PR	05-AUG-1999;	99US-0147260P.								
PR	06-AUG-1999;	99US-0147303P.								
PR	06-AUG-1999;	99US-0147416P.								
PR	09-AUG-1999;	99US-0147493P.								
PR	09-AUG-1999;	99US-0147935P.								
PR	10-AUG-1999;	99US-0148171P.								
PR	11-AUG-1999;	99US-0148319P.								
PR	12-AUG-1999;	99US-0148341P.								
PR	13-AUG-1999;	99US-0148565P.								
PR	13-AUG-1999;	99US-0148684P.								
PR	16-AUG-1999;	99US-0149368P.								
PR	17-AUG-1999;	99US-0149175P.								
PR	18-AUG-1999;	99US-0149426P.								
PR	20-AUG-1999;	99US-0149722P.								
PR	20-AUG-1999;	99US-0149723P.								
PR	20-AUG-1999;	99US-0149929P.								
PR	23-AUG-1999;	99US-0149902P.								
PR	23-AUG-1999;	99US-0149930P.								
PR	25-AUG-1999;	99US-0150566P.								
PR	26-AUG-1999;	99US-0150884P.								
PR	27-AUG-1999;	99US-0151065P.								
PR	27-AUG-1999;	99US-0151066P.								
PR	27-AUG-1999;	99US-0151080P.								
PR	30-AUG-1999;	99US-0151303P.								
PR	31-AUG-1999;	99US-0151438P.								
PR	01-SEP-1999;	99US-0151930P.								
PR	07-SEP-1999;	99US-0152363P.								
PR	10-SEP-1999;	99US-0153070P.								
QY	89	GAGTATACATGACAAACGGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCACACTGTTG 148	QY	89	GAGTATACATGACAAACGGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCACACTGTTG 148					
Db	191	GACTTTACTACGAAAAAGCGGTAAAGCACATGGCTCGAAACCTTGGTTTCAGCTTGTAG 250	Db	191	GACTTTACTACGAAAAAGCGGTAAAGCACATGGCTCGAAACCTTGGTTTCAGCTTGTAG 250					
QY	149	GCTTTCTGTGCTACTTCCATATATATCTGTGCAATTTAAACACATGCAACACTGATA 208	QY	149	GCTTTCTGTGCTACTTCCATATATATCTGTGCAATTTAAACACATGCAACACTGATA 208					
Db	251	GGTTTCTTTAAACCTTCTTGTCTATTATTACTTAAAGCCTGAGCCGCTCAAGACTTAAA 310	Db	251	GGTTTCTTTAAACCTTCTTGTCTATTATTACTTAAAGCCTGAGCCGCTCAAGACTTAAA 310					
QY	209	GAGATGGAAGAAAGAACCTCACCTAG-----GAAACCGTGTATTGGTTTACGTAGTCTTG 262	QY	209	GAGATGGAAGAAAGAACCTCACCTAG-----GAAACCGTGTATTGGTTTACGTAGTCTTG 262					
Db	311	CCATTACCAAAAAAAACTACTTCTTCTTGGACATCTCTTTAGTATATATTGACTTG 370	Db	311	CCATTACCAAAAAAAACTACTTCTTCTTGGACATCTCTTTAGTATATATTGACTTG 370					
QY	263	GACTTCTTTGTAGGACGACATTCATCTGTACTCCATTGGACTTCTTTACTTTACCGTTT 322	QY	263	GACTTCTTTGTAGGACGACATTCATCTGTACTCCATTGGACTTCTTTACTTTACCGTTT 322					
Db	371	GCTTGTCTTGTGCTGGACATTTTGTACTCATTTTGGGCTACTTTTACCTTCTCTGTCT 430	Db	371	GCTTGTCTTGTGCTGGACATTTTGTACTCATTTTGGGCTACTTTTACCTTCTCTGTCT 430					
QY	323	CTACCTATTCCTGTGATCTGTGATCTCAGTTAGCCTTCAAGCTTCTTCTTCTTATTTTC 382	QY	323	CTACCTATTCCTGTGATCTGTGATCTCAGTTAGCCTTCAAGCTTCTTCTTCTTATTTTC 382					
Db	431	CAACTTTCTTTTGTATCTCTGCTGCAATTTGGCTTTTAAACCGCTCTTCTCTTACTTCC 490	Db	431	CAACTTTCTTTTGTATCTCTGCTGCAATTTGGCTTTTAAACCGCTCTTCTCTTACTTCC 490					
QY	383	TTAACTCAGAAAAACTTACCCCTATCATTTTAAATTTCTTTTCTTCTTAACTATATCTT 442	QY	383	TTAACTCAGAAAAACTTACCCCTATCATTTTAAATTTCTTTTCTTCTTAACTATATCTT 442					
Db	491	TAACTCAGAAAAAATCACCAATTTACTCAATTCATCTTCTTCTTAACTATCTT 550	Db	491	TAACTCAGAAAAAATCACCAATTTACTCAATTCATCTTCTTCTTAACTATCTT 550					
QY	443	CCACCTTACTTGTCAATTCAATATGAGGAGACAGACTCCACAAAAGTTACAAAAGGAG--- 499	QY	443	CCACCTTACTTGTCAATTCAATATGAGGAGACAGACTCCACAAAAGTTACAAAAGGAG--- 499					

QY 812 AGCTCTCTCTCTATCTCAATGCAATAGCGTTTGGACATCCAGTGGTTCCTATCT 871
Db 68018 TGGTGCTTCGCTGTTTCAACCTTATTAGTAGCTCTCACTCATTTGTACGCTCTCG 68077
QY 872 TGCTGTGAATCATTTTCCATGACAAATGAATGGTTAAAGTGATTCTCTATGATCCTAG 931
Db 68078 CGGCCATTGGGTGTTCCATGACAGCTGACTGAGGTTAGATGGTGGCATGCCATCG 68137
QY 932 CTATTTGGGGTTTCACTTCTCTATGTCTACCAACAATATCTTGATGACAAAACTTGAAGA 991
Db 68138 CCTTCACAGGATTCACGCTTTTATATCTACCAAGTACTCTTGATGACTTGAAAGTACAA 68197
QY 992 AAATCATGNA 1002
Db 68198 GAGCAGAGAA 68208
RESULT 13
AAC37652
ID AAC37652 standard; DNA; 1403 BP.
XX
AC AAC37652;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 18172.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144003P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.

PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-015138P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-015753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	18-OCT-1999;	99US-0159638P.
PR	21-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query Match

Best Local Similarity

Matches

20.9%;

Score

225.8;

DB

3;

Length

1403;

54.0%;

Pred. No. i.le-53;

0;

Mismatches

407;

Indels

12;

Gaps

1;

Qy	92	TATACTATGACAAACGGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCAACTTCTGGCT	151
Db	345	TTTATTACAAGAGGTGGAAACAGTAAATGGATCTCTACTCTTGTCCAACTTGTGGTT	404
Qy	152	TTCCGTGTCTACTTCCATATTATATCTTGTCAATTAACACACATGCAACCACTGATAGAG	211
Db	405	TTCCGATTCTCTATCTCCCTCTTGTTCCTTCCCTGCTTCCACACTCTCTTCTCATCTTCT	464
Qy	212	ATGGAACAAACCTACCTACCTAGGAACCGTGTATTTGGTTTAGGTAGTGTGGACTTCTTG	271
Db	465	GT-----TCTTCAAGACTCTGGTTTGGATTATCTTCTCTGGTTTCTCTA	512
Qy	272	TAGGAGCAGATTGCTATCTCTACTCTCCATTGGACTTCTTTACTTTACCCGTTTCTACCTATT	331
Db	513	TTGGTTTAGACAATCTTTTATATCTTTTGGTCTTTTGTATCTCTCTGCTTCACTTATT	572
Qy	332	CCCTGATCTGTGCACTCTCAGTTAGCCTTCAATGCTTCTTCTTCTTATTTTCTTAACCTAC	391
Db	573	CGATTCTTTGTTCTTACAGTTAGCTTTCATGCTGCTCTCTCTTATTACATCAATCTCTC	632
Qy	392	AAAACTTACCCCTATCAATTTTAAATCTCTTTTCCCTCTTAACTATATCTTCCACCCTAC	451
Db	633	AGAAATCACTTGTGTGATTCTCTCTCTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT	692
Qy	452	TTGCATTCAATAATGAGGAGACAGACTCCACAAAGTTACAAAGGAGAGTATCTCAAAG	511
Db	693	TTTCTCTTGAGATGATTCCAATAGCCCATCAGGAGATTCTAAGTGGAGTTACTTGTATTG	752
Qy	512	GTTTCATATGACCGTTTGTGCGTCTGCTGGTTAGTCTTAGTCTTAGTCTTATCCCTACAAACG	571
Db	753	GTTGTTTATGTACTGTTTTTGTCTTCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	812
Qy	572	TAGCTTTCTAAAGTCTTAAAGAACAAATTTCTCAGAAAGTTATCGATGATGATATCT	631
Db	813	TTTCTTTTCGAGAAATGTTCTCAAGAGTGAGAGCTTCTCTATGGTTCTCGAGATGCAAACT	872
Qy	632	ACGTGAGTCTAGTGGCGAGTTGTGTAGCGTGGTGGGCTTTTGTGTAGCAGTGAAGTGA	691
Db	873	ATAGTCTGCTTGTGGCTTCTTGTGTGGGTTATCGGGTTGTTCGAGCGGGGAATGGA	932
Qy	692	AACTTTGAGCAGTGAATAAGATAACTACAAACATGGAAGGATCTCTACATTAAGAAC	751
Db	933	TGTTGTTGAGTGTGGAGATGGAAGAGTTTCAAGAGGTCAGTCAATTTATGTTTGTACTT	992
Qy	752	TAGTGTGACAGCTGTTTACCTGGCAGTTATCTCCATCGGTGGCACAGGACTGATCTTCG	811
Db	993	TGGTGGGACCGCGGTTTCGTGGCAATTTGGGTTCTGTAGGAGCGGTGGCGCTTATATTT	1052
Qy	812	AGCTCTCTCTCTATTCTCAAATGCAATAAGCGTTTTTGGGACTCCCACTGGTCTCTATCT	871
Db	1053	TGGTTCTTCTGCTGTTTTTCGNACTTATTTGTTGATGATGATGATGATGATGATGATGAT	1112
Qy	872	TGGCTGTAATCAATTTTCCATGACAAAATGAATGAATGAATGAATGAATGAATGAATGA	931
Db	1113	CAGCCATTGGCGTGTTCACGACAAAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	1172
Qy	932	CTATTGGGGTTTCACTTCTCTACCAACAAATATCTTGATGATGATGATGATGATGATGAT	991
Db	1173	CTTTCATGGGGTTTCGGGTTTTTATATCTACCAGAAATTTATCTTGTGATGATGATGAT	1232
Qy	992	AAATCATGAA	1002
Db	1233	GAGCAGGAA	1243

RESULT 14

ADA70748

ID

ADA70748

standard; DNA;

2175

BP.

CC illustrate the invention.

```
XX
SQ Sequence 1152 BP; 294 A; 260 C; 248 G; 350 T; 0 U; 0 Other;

Query Match      19.6%; Score 211.6; DB 8; Length 1152;
Best Local Similarity 57.7%; Pred. No. 1.2e-49;
Matches 404; Conservative 0; Mismatches 284; Indels 12; Gaps 1;

QY 92 TATACATGACACGAGGAGAACAGTAAATGGCTAGCAACGGTAGTTCAACTTGTGGCT 151
   |||||
Db 329 TCTACTACAATCAAGCGGCAATAGCAAGTGGATGCCACATTCGCCAAACTGCTGGCT 388

QY 152 TTCC-----TGTGCTACTTCCATATATATCTTGTCAATTTAAAAACATGCAA 199
   |||||
Db 389 TTCCGATTTGTTCATTCGCCATTTCTTTCCATTCAAAGACATCTTCTACACAACAG 448

QY 200 CAACGTAGAGATGGAAGAAAGAACCTCACCTAGGAACCGTGTATTTGGTTTACGTAGGC 259
   |||||
Db 449 TCACTAGTAGTCTCGCCCTACAATTTCTATCCCAAAATTAATCTGATATATGTTGCC 508

QY 260 TTGGACTTCTTTGAGGACAGATTGCTATCTGTACTCCATTGGACTTCTTACTTACCG 319
   |||||
Db 509 TGGGCTCATCATGCTCGACACGACTTGATGTATTCCTATGGCTACTATATCTTCCGG 568

QY 320 TTCTTACCTATTCCTGTATCTGTGCATCTCAGTTAGCCTTCAATGCTTTCTTCTCTTAT 379
   |||||
Db 569 TCTCAACATATTCGCTCATCTGTGCTAGTCACTGCTTGCCTTCAATGCTGTCTTCTCATAT 628

QY 380 TTCTTAACTCACAAAACCTTACCCCTATCATTTTAAATTTCTTTTCTCTTAACATATAT 439
   |||||
Db 629 TCCTCAATGCTCAAAAATTCACCCCTCTGATTTTCAATTCGGTAGTCTCTCTTACGTTTT 688

QY 440 CTTCCACCTACTTGCATTCATATATAGGAGACAGACTCCACAAAAGTTACAAAAGGAG 499
   |||||
Db 689 CTGCTTCACTCTTGGAGTTGATGAAGATCTCAGGGAACTACTAGTATATCACATGGGA 748

QY 500 AGTATGTCAAAGGTTTCATATGCAACCGTTGCTGCTGTGCTGTATGGTCTAGTCTTAT 559
   |||||
Db 749 AGTACATTTTGGGTTTCTCTGTGACACTAGGGGCATCAGCTACATCTCGCTCATCTCTCT 808

QY 560 CCCTACAAAGCTAGCCTTTCTAAAGTCTTAAAGAGCAAAATTTCTCAGAAAGTTATGG 619
   |||||
Db 809 CCCTGATGCAAGTCACATTTGAGAAGGTTATTAAGAGGGAGACCTTCTCAGTTGTGTGA 868

QY 620 ATATGATAATCTACGTAGTCTAGTGGCCAGTTGTGTTAGCGTGGTGGGCTTTTGGCTA 679
   |||||
Db 869 ACATCCAGATATATACAGCTCTCGTGGCAACATTTGGCTTCTCTTGTGGGTTATTTGCA 928

QY 680 GCAGTGAGTGGAAAACCTTTGAGCAGTGAATGGATACTACAAAACATGGGAAGGTATCCT 739
   |||||
Db 929 GTGGTGAATGGATGACTTTACAAGGAGAGATGCATTCAGTCTCGGAAGCTGTICAT 988

QY 740 ACATTATGAACCTAGTGTGGACAGCTGTTACTTGGCAGTT 779
   |||||
Db 989 ATGTAATGACACTGTGTGGACGGCTATATCTTGGCAGAT 1028
```

Search completed: November 1, 2004, 07:20:05
Job time : 571.648 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:48:03 ; Search time 119.284 Seconds
(without alignments)
7704.726 Million cell updates/sec

Title: US-09-913-767-4

Perfect score: 1293

Sequence: 1 ggaagctctcttagagtggt.....tatgatcaaacatatattcc 1293

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/1/ina/5A-COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B-COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A-COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B-COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCTUS-COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79.2	6.1	7218	1	US-08-232-463-14
2	44.2	3.4	1141	4	US-09-806-708B-22
3	42.8	3.3	399	4	US-09-621-976-8976
4	41.4	3.2	1141	4	US-09-806-708B-22
5	38.2	3.0	696	3	US-09-461-697-193
6	38.2	3.0	699	3	US-09-461-697-191
7	38.2	3.0	717	3	US-09-461-697-189
8	38.2	3.0	774	3	US-09-461-697-187
9	38.2	3.0	819	3	US-09-461-697-185
10	38.2	3.0	1669	3	US-09-461-697-184
11	38.2	3.0	1664976	4	US-08-916-421B-1
12	38.2	3.0	1664976	4	US-09-692-570-1
13	38.2	2.9	11917	4	US-09-566-921-32
14	35.6	2.8	264	4	US-09-107-532A-2112
15	35.6	2.8	600	4	US-09-107-532A-3027
16	35.6	2.8	832	4	US-09-621-976-2813
17	35.6	2.8	1047	3	US-09-198-955A-7
18	35.6	2.8	1047	3	US-09-694-531-7
19	35.6	2.8	1047	4	US-10-072-152-7
20	35.6	2.8	1266	4	US-09-248-796A-6306
21	35.2	2.7	504	3	US-09-134-001C-2504
22	35.2	2.7	1827	4	US-09-270-767-1308
23	35.2	2.7	1827	4	US-09-270-767-16590
24	35	2.7	312	4	US-09-621-976-18100
25	34.6	2.7	1394	3	US-09-247-155-76
26	34.4	2.7	474	4	US-09-621-976-18033
27	34.2	2.6	289	3	US-09-007-005-17

c	28	34.2	2.6	289	3	US-09-244-796-17	Sequence 17, Appl
	29	34.2	2.6	832	4	US-09-621-976-2813	Sequence 2813, Ap
	30	34.2	2.6	246240	2	US-08-724-394A-20	Sequence 20, Appl
	31	34.2	2.6	246240	2	US-08-724-394A-21	Sequence 21, Appl
	32	34.2	2.6	246240	2	US-08-724-394A-21	Sequence 22, Appl
	33	34	2.6	1794	4	US-09-248-796A-2292	Sequence 2292, Ap
c	34	34	2.6	1923	4	US-09-134-000C-466	Sequence 466, App
	35	33.8	2.6	1332	4	US-09-642-894A-1	Sequence 1, Appli
	36	33.6	2.6	2072	4	US-09-774-528-107	Sequence 107, App
	37	33.6	2.6	2440	4	US-09-774-528-106	Sequence 106, App
	38	33.6	2.6	80246	3	US-09-078-294-4	Sequence 4, Appli
	39	33.6	2.6	80595	3	US-09-078-294-3	Sequence 3, Appli
	40	33.4	2.6	16550	4	US-08-916-421B-3	Sequence 3, Appli
	41	33.4	2.6	16550	4	US-09-692-570-3	Sequence 3, Appli
c	42	33.2	2.6	231	3	US-09-461-697-207	Sequence 207, App
	43	33.2	2.6	282	3	US-09-461-697-205	Sequence 205, App
c	44	33.2	2.6	306	3	US-09-461-697-203	Sequence 203, App
	45	33	2.6	1473	4	US-09-244-805-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Hardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ59pt-F15
US-08-232-463-14

Query Match 6.1%; Score 79.2; DB 1; Length 7218;

[illegible]

```
Qy 670 AAC 672
Db 315 ATC 313

RESULT 6
US-09-461-697-191/c
; Sequence 191, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 191
; LENGTH: 699
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-191

Query Match 3.0%; Score 38.2; DB 3; Length 699;
Best Local Similarity 47.3%; Pred. No. 0.08;
Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 430 TTCTCTTCTTCCACCAATCTTGGATCAGTTTACATCGTTACTGGACATATTAGTCTGCT 489
Db 558 TTCTCTCTCTTTTCACTCTCTTTGACATCTTTTACATCTTTCCCTCTTTTATCTCCCTT 499

Qy 490 AACTCTTATATGTCCTCTGTGGTTTACTATACCACTTACTTACCACTTCTCTCTCCCTCATC 549
Db 498 CTCCTTTCCACTTCACTCTCTTTTCCAACTCCTGTTCTTTCTGTCTCTCTCTTT 439

Qy 550 TTGGCCTCACAAATGGCCCTTCACTGCTTCTTCACTATTTTCTAACTCGCAGAGTTC 609
Db 438 TTCACTCTCTCTCTCTCTTTTCCATCTTCTCCATTTCTCTTCCGTTTCACTCTTTT 379

Qy 610 ACACCTTTCAATGGAATCTCTGTTTCTCTCTTACTATTTCTCTGCTCTGCTCTGCTGTC 669
Db 378 TCCATCTCTTCTCTTTTGGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTCC 319

Qy 670 AAC 672
Db 318 ATC 316

RESULT 7
US-09-461-697-189/c
; Sequence 189, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
```

```
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 189
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-189

Query Match 3.0%; Score 38.2; DB 3; Length 717;
Best Local Similarity 47.3%; Pred. No. 0.081;
Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 430 TTCTCTTCTTCCACCAATCTTGGATCAGTTTACATCGTTACTGGACATATTAGTCTGCT 489
Db 576 TTCTCTCTCTTTTCACTCTCTTTGACATCTTTTACATCTTTCCCTCTTTTATCTCCCTT 517

Qy 490 AACTCTTATATGTCCTCTGTGGTTTACTATACCACTTACTTACCACTTCTCTCTCCCTCATC 549
Db 516 CTCCTTTCCACTTCACTCTCTTTTCCAACTCCTGTTCTTTCTGTCTCTCTCTCTTTT 457

Qy 550 TTGGCCTCACAAATGGCCCTTCACTGCTTCTCTCTCATATTTTCTAACTCGCAGAGTTC 609
Db 456 TTCACTCTCTCTCTCTCTTTTCCATCTTCTCCATTTCTCCGTTTCACTCTTTT 397

Qy 610 ACACCTTTCAATGGAATCTCTGTTTCTCTCTTACTATTTCTCTGCTCTGCTGTC 669
Db 396 TCCATCTCTCTCTCTCTTTTGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTCC 337

Qy 670 AAC 672
Db 336 ATC 334

RESULT 8
US-09-461-697-187/c
; Sequence 187, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 187
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-187

Query Match 3.0%; Score 38.2; DB 3; Length 774;
Best Local Similarity 47.3%; Pred. No. 0.085;
Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 430 TTCTCTTCTTCCACCAATCTTGGATCAGTTTACATCGTTACTGGACATATTAGTCTGCT 489
Db 633 TTCTCTCTCTTTTCACTCTCTTTGACATCTTTTACATCTTTCCCTCTTTTATCTCCCTT 574

Qy 490 AACTCTTATATGTCCTCTGTGGTTTACTATACCACTTACTTACCACTTCTCTCTCCCTCATC 549
Db 573 CTCCTTTCCACTCTCACTCTCTTTTCCAACTCCTGTTCTTTCTCTCTCTCTCTCTTTTCC 514
```



```

;
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 2112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...264
; SEQUENCE DESCRIPTION: SEQ ID NO: 2112:
US-09-107-532A-2112

Query Match 2.8%; Score 35.6; DB 4; Length 264;
Best Local Similarity 60.2%; Pred. No. 0.29;
Matches 59; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1044 TTCCATAACTGCTGTGGGATTGCCTATAGTTCCAGTTGTAGCAGTGATTGTTTCCATGA 1103
Db 123 TTTTAAGACTGTAGCGGAGTCTGCTTTTCTGTTGTTTAGAGATTGTTCCAAAAGT 64

QY 1104 TAAATGAACGGCTCAAGATCTTCCATCATTTTAG 1141
Db 63 GGAATGATCTCTCACATTCTTTTCACTGATTTCAG 26

RESULT 15
US-09-107-532A-3027
; Sequence 3027, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3027:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...600
; SEQUENCE DESCRIPTION: SEQ ID NO: 3027:
US-09-107-532A-3027

Query Match 2.8%; Score 35.6; DB 4; Length 600;
Best Local Similarity 60.2%; Pred. No. 0.49;
Matches 59; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1044 TTCCATAACTGCTGTGGGATTGCCTATAGTTCCAGTTGTAGCAGTGATTGTTTCCATGA 1103
Db 234 TTTTAAGACTGTAGCGGAGTCTGCTTTTCTGTTGTTTAGAGATTGTTCCAAAAGT 293

QY 1104 TAAATGAACGGCTCAAGATCTTCCATCATTTTAG 1141
Db 294 GGAATGATCTCTCACATTCTTTTCACTGATTTCAG 331

Search completed: November 1, 2004, 22:40:49
Job time : 125.284 secs
```

THIS PAGE BLANK (USPTO)

Qy	22	ATGAAATAACTCAAGTAATCTATGTCAATGGTAAGCAAGATGCATCTCGTAGTGTAGAT	81
Dd	1	ATGAAATAACTCAAGTAATCTATGTCAATGGTAAGCAAGATGCATCTCGTAGTGTAGAT	60

Qy 82 TACTTGATTCTTTTCGCTAAACCTGTTGTTTGTATGTTTTCAGATCATATAACAACATAGAA 141
Db 61 TACTTGATTCTTTTCGCTAAACCTGTTGTTTGTATGTTTTCAGATCATATAACAACATAGAA 120
Qy 142 GCAAACTTAAACAGGTGAGGAGAAATGAATACCACTGGAATCGAATCTTCCTCGGTA 201
Db 121 GCAAACTTAAACAGGTGAGGAGAAATGAATACCACTGGAATCGAATCTTCCTCGGTA 180
Qy 202 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTTCCATTTACGTTCTTTGTCCCTT 261
Db 181 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTTCCATTTACGTTCTTTGTCCCTT 227
Qy 262 GCTTGCCCAAGCACATTTCTACAAATTTGGCAGAGTTTACTATGAATAATGGTGGGAAGGT 321
Db 228 -----AGTTTACTATGAATAATGGTGGGAAGGT 255
Qy 322 ACATGGATGGAAACACTTGTCCAACTAATCGGCTTCCCTGTTCTGTTTCTCTCCGCTTC 381
Db 256 ACATGGATGGAAACACTTGTCCAACTAATCGGCTTCCCTGTTCTGTTTCTCTCCGCTTC 315
Qy 382 TTTTCCCAAAACCAAAATCCAAACCAAGAGAGATTTTCAAGAAAGTTCTCTTCCTTC 441
Db 316 TTTTCCCAAAACCAAAATCCAAACCAAGAGAGATTTTCAAGAAAGTTCTCTTCCTTC 375
Qy 442 ACCATTCTTGATCAGTTTACATCGTTACTGGACTATATAGTGTCTGCTAACTCTATATG 501
Db 376 ACCATTCTTGATCAGTTTACATCGTTACTGGACTATATAGTGTCTGCTAACTCTATATG 435
Qy 502 TCCTCTGTTGGTTACTACTTACCAAGTTTCTACTTTCCTCCCTCATCTTGGCCTCAAA 561
Db 436 TCCTCTGTTGGTTACTACTTACCAAGTTTCTACTTTCCTCCCTCATCTTGGCCTCAAA 495
Qy 562 TTGGCCTTCACTGCTTCTTCTCATATATTTCTAAACCTGCGAAGATTCACACCTTTCAAT 621
Db 496 TTGGCCTTCACTGCTTCTTCTCATATATTTCTAAACCTGCGAAGATTCACACCTTTCAAT 555
Qy 622 GTGAATCTCTGTTTCTCTTACTATTTTCTGCTCCCTCCCTGCTGCTGCTGCTGCTGCTG 681
Db 556 GTGAATCTCTGTTTCTCTTACTATTTTCTGCTCCCTCCCTGCTGCTGCTGCTGCTGCTG 615
Qy 682 GAAACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGGAATAATATGATACCAATGGT 741
Db 616 GAAACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGGAATAATATGATACCAATGGT 675
Qy 742 GCTTCTGCTGGATGATGCTGCTATCCCTGGTGAACACTGATCCTCAGGAAGGTTTAA 801
Db 676 GCTTCTGCTGGATGATGCTGCTATCCCTGGTGAACACTGATCCTCAGGAAGGTTTAA 735
Qy 802 AAGAGCAAAACATTTCTCAACGGTCACTGACTTGGTGGCTTACCCTTCTAGTTGCAAGC 861
Db 736 AAGAGCAAAACATTTCTCAACGGTCACTGACTTGGTGGCTTACCCTTCTAGTTGCAAGC 795
Qy 862 TGTGTGGTTCATAGGACTTTTCGAAGCGGGAGTGGAAAACCTTTTAAACAGTGAGATG 921
Db 796 TGTGTGGTTCATAGGACTTTTCGAAGCGGGAGTGGAAAACCTTTTAAACAGTGAGATG 855
Qy 922 GAAACATCAAACTGGGGAAGTGCCATAGCTTTATGACTTTTGGCTCGATAGCTATTTCC 981
Db 856 GAAACATCAAACTGGGGAAGTGCCATAGCTTTATGACTTTTGGCTCGATAGCTATTTCC 915
Qy 982 TGGCAAGTCTACACATTTGGGCTGCTGGAGCTGATCTTTGATGCTATCTTCTGTTCTCC 1041
Db 916 TGGCAAGTCTACACATTTGGGCTGCTGGAGCTGATCTTTGATGCTATCTTCTGTTCTCC 975
Qy 1042 AATTCCATAACTGCTGTGGGATTTGCTATAGTTTCCAGTTGTAGCAGTGATTTGTTTCCAT 1101
Db 976 AATTCCATAACTGCTGTGGGATTTGCTATAGTTTCCAGTTGTAGCAGTGATTTGTTTCCAT 1035
Qy 1102 GATAAAATGAACGGCTCAAAAGATCTTCTCCATCATTTTATAGCTATCTGGGGATTCATTTCA 1161
Db 1036 GATAAAATGAACGGCTCAAAAGATCTTCTCCATCATTTTATAGCTATCTGGGGATTCATTTCA 1095
Qy 1162 TTTGTCTATCAGCACTACTCTCGAGAAAGATTTGAAGACTTAGCCACACAAGTCCCTGTA 1221

Db 1096 TTTGCTATCAGCACTACCTCGAGAAAGAAAGTTGAGAGCTAGCCACACAAGTCCCTGTA 1155
Qy 1222 GGAGATCCTCATCTACTACTCCTGCTGAGGAAGGTGCACACAAACATACATAGTG 1273
Db 1156 GGAGATCCTCATCTACTACTCCTGCTGAGGAAGGTGCACACAAACATACATAGTG 1207
RESULT 2
US-09-938-842A-2212
; Sequence 2212, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2212
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2212
Query Match 89.1%; Score 1152; DB 11; Length 3387;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1207; Conservative 0; Mismatches 0; Indels 45; Gaps 1;
Qy 22 ATGGAATAACTCAAGTAATCTATGTCAATGGTAAGCAAGATGCAATCTCGTAGTGTAGAT 81
Db 1 ATGGAATAACTCAAGTAATCTATGTCAATGGTAAGCAAGATGCAATCTCGTAGTGTAGAT 60
Qy 82 TACTTGATTCTTTTCGCTTAACCTGTTGTTTTCAGATCATATAACAACATAGAA 141
Db 61 TACTTGATTCTTTTCGCTTAACCTGTTGTTTTCAGATCATATAACAACATAGAA 120
Qy 142 GCAAACTTAAACAGGTGAGGAGAAATGAATACCACTGGAATAATCGAATCTTCCTCGGTA 201
Db 121 GCAAACTTAAACAGGTGAGGAGAAATGAATACCACTGGAATAATCGAATCTTCCTCGGTA 180
Qy 202 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTTCCATTTACGTTCTTTGTCCCTT 261
Db 181 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTTCCATTTACGTTCTTTGTCCCTT 227
Qy 262 GCTTGCCAAGCACTTTCCTACAAATTTGGGCGAGAGTTTACTATGAATAATGGTGGGAAGGT 321
Db 228 -----AGTTTACTATGAATAATGGTGGGAAGGT 255
Qy 322 ACATGGATGGAAACACTTGTCCAACTAATCGGCTTCCCTGTTCTGTTTCTCTCCGCTTC 381
Db 256 ACATGGATGGAAACACTTGTCCAACTAATCGGCTTCCCTGTTCTGTTTCTCTCCGCTTC 315
Qy 382 TTTTCCCAAAACCAAAATCCAAACCAAGAGAGATTTTCAAGAAAGTTCTCTTCCTTC 441
Db 316 TTTTCCCAAAACCAAAATCCAAACCAAGAGAGATTTTCAAGAAAGTTCTCTTCCTTC 375
Qy 442 ACCATTCTTGATCAGTTTACATCGTTACTGGACTATATAGTGTCTGCTAACTCTATATG 501
Db 376 ACCATTCTTGATCAGTTTACATCGTTACTGGACTATATAGTGTCTGCTAACTCTATATG 435
Qy 502 TCCTCTGTTGGTTACTACTTACCAAGTTTCTACTTTCCTCCCTCATCTTGGCCTCAAA 561

Qy 1210 ACAAGTCTGTAGGATCCTCATCTACTACTGCTGAGGAGGTCACACA 1260
Db 1139 ACTACTAGTCATGTTCCAAAGACTTCTTCCACCTATAGAAGAGGTTTCATAGA 1189

RESULT 4

US-09-770-152-44

; Sequence 44, Application US/09770152

; Publication No. US20020040489A1

; GENERAL INFORMATION:

; APPLICANT: Gorlach, Jorn

; APPLICANT: An, Yong-Qiang

; APPLICANT: Hamilton, Carol M.

; APPLICANT: Price, Jennifer L.

; APPLICANT: Raines, Tracy M.

; APPLICANT: Yu, Yang

; APPLICANT: Rameaka, Joshua G.

; APPLICANT: Page, Amy

; APPLICANT: Matthew, Abraham V.

; APPLICANT: Ledford, Brooke L.

; APPLICANT: Woessner, Jeffrey P.

; APPLICANT: Haas, William David

; APPLICANT: Garcia, Carlos A.

; APPLICANT: Kriker, Maja

; APPLICANT: Slader, Ted

; APPLICANT: Davis, Keith R.

; APPLICANT: Allen, Keith

; APPLICANT: Hoffman, Neil

; APPLICANT: Hurban, Patrick

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis

; FILE REFERENCE: 2025US (PARA-014PRV)

; CURRENT APPLICATION NUMBER: US/09/770,152

; PRIOR FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: 60/178,503

; PRIOR FILING DATE: 2000-01-27

; NUMBER OF SEQ ID NOS: 999

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 44

; LENGTH: 592

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-770-152-44

Query Match 27.4%; Score 354.4; DB 9; Length 592;

Best Local Similarity 81.8%; Pred. No. 2.2e-94;

Matches 409; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 180 GGAATCGAATCTTCGTCGGTACCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTC 239
Db 93 GGAACCTGAATCATTTTCAGTACCTCAACGGAAGAACTGTAAAGGTGGCTCCGTGCTC 152
Qy 240 CATTTACGTGTTCTTTGTCTTGTCTTGTTCGCAAGCACTTTCTACAAATTTTGGGCAGAGTTTA 299
Db 153 CATATACCAATCTTTGTCTCTTCGCAACCACTTGTCTACAGTTCTGGGTAGACTGTA 212
Qy 300 CTATGAAATGGTGGGAAGAGTACATGATGGGAACACTTGTCCAACTAATTCGGCTTCCC 359
Db 213 CTATGAAATGGGGAAGACATATGTGTAACTTCTTCAACTCATTTGGCTTCCC 272
Qy 360 TGTCTGTCTCTCTCTCGCTTCTTTTCCCAACCAAAATCCCAACCAAGAGCAGA 419
Db 273 TGTACTGATCTCTCTCGCTTCTTTCTCGAATCAGGCAACCAAAATCAACAGATACAAA 332
Qy 420 TTTCAGAAAGTTCTCTCTTCCACCAATCTTTGGATCAGTTTACATCGTTACTGGCAATTT 479
Db 333 TTTCACTCAGTCCCTTCTTCCACCACTTGTGATCGGTTTACTTGTGCACTGGCTGCT 392
Qy 480 AGTGTCTCTAACTCTTATATGTCCTCTGTGGTTTACTATACCTTACAGTTTCTACTTT 539
Db 393 AGTGTCTCTTATGCTTATTTGTCGCAAGTTGGGTTGCTTTTACTTACCACTCTCTACTTT 452
Qy 540 CTCCTCATCTTGGCCTCACAAATTGGCCTTCACTGCCTTCTCTCATATTTTCTAAACCTC 599

Db 453 CTCCCTCATCTTGGCCTCACAGTTGGCTTCACTGCCTTTTCTCATATTTCCTTAACTC 512
Qy 600 GCAGAAAGTTCCACACCTTTTCATTGTGAATTCCTGTGTTTCTCTTACTATTTCTCTGCGCT 659
Db 513 GCAAAAGTTCACTCCCTTTGATAGTCAATTCCTTGTGTTTCTCTTACGGTTTCTCTGCGCT 572
Qy 660 CCTCGTGGTCAACACTGATT 679
Db 573 CCTCGTGGTCAACACTGATT 592

RESULT 5

US-10-424-599-102519

; Sequence 102519, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 102519

; LENGTH: 2214

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE: INFORMATION: Clone ID: PAT_MRT3847_63592C.1

; US-10-424-599-102519

Query Match 23.8%; Score 307.6; DB 16; Length 2214;
Best Local Similarity 58.5%; Pred. No. 4e-80;
Matches 580; Conservative 0; Mismatches 399; Indels 13; Gaps 2;

Qy 204 TCAATCGAAGAACTATAAGAAATGGCTTCGTATTTCCTATTATTCGTGTTCTTTGCTTGC 263
Db 389 TAAATATAAGCATGGCAATGGTGTCTTTGTGACACTCAGCATAGCTTTCTTATAGT 448
Qy 264 TTGCCAAGCACTTTCTACAAATTTGGGCAGAGTTTACTATGAAATCGTGGAGAGTAC 323
Db 449 TGGCAATCTGCTGCTGTTATCTTGGAGATTTTATATGATCAGGGTGAATAGTAA 508
Qy 324 ATGGATGGGAACACCTTGTCCAACTAATCGGCTTCCCTGTTCTGTTTCTTCTTCGCTTCTT 383
Db 509 ATGGATGGCTACTCTAGTTCAAACTGCTGCTTCCCGATCTTG-----TTTCA 556
Qy 384 TTCCAAACCAAAATCCCAACCAACAGAGAGATTTTCAGAAAGTTCTTCTTCCCTCAC 443
Db 557 TCCATATTTTACAATTCCTTCACTCCAGAGGCTTCAACTTCTGCTTTCACCTCCCACTCAA 616
Qy 444 CATCTTTGGATCAGTTTACATCGTTAC-TGGACTATTAGTGTCTGCTAACTCTTATATGT 502
Db 617 AATTATTTCTTTGATATATTTTGGTCTTTGGAGTCTTAAATGCTGCTGCAATATGATGT 676
Qy 503 CCTCTGTGTTTACTATATCTTACCAGTTTCTACTTCTCCCTCATCTTGGCCTTCACAAT 562
Db 677 ACTCCACTGGACTCTTATACCTCTCGGCTTCTACCTATTTCGCTGATTTGTGCATCAGT 736
Qy 563 TGGCCTTCACTGCGCTTCTTCTCATATTTTCTAACTCGCAGAGTTTTCACACCTTTCATTG 622
Db 737 TAGCTTTTAAATGCAAGTTTCTCTCATATTTTATCAATTTCTCAAAAGTTTCACTGCGCTTGA 796
Qy 623 TGAATCTCTGTTTCTCTTACTATTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 682
Db 797 TAAACTCTACAGTGGTGTCTCACCTTTTATCTGCTGCACTTCTTGTGTTTAAAGACACAG 856
Qy 683 AAAACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGATAAATATGTACCAATTGGTG 742

Db 857 ATGAACCATCTGGTTCTTCCAAAGGGAAGTACATATATGGTTTCCATGTACCCCTTGGAG 916
Qy 743 CTTCTGCTGGGATGGATTGCTCTATCCCTGTGTACAACTGATCCTCAGGAAGGTTTTAA 802
Db 917 CTTCTGAGTGTAATCTCTTTTGGTTTCCCTCATGAGCTGACCTTTGAGAGGTTCTGA 976
Qy 803 AGAAGCAAAATTTCTCAAGGTCACCTGATGCTGTGCTGTACCAATCTCTAGTTGCAAGCT 862
Db 977 AGAAGGAAACATTTCTGTGTTTGGAAATGCAAAATCTACACATCATCTGTTGCTCTG 1036
Qy 863 GTGTGTTCTCATAGGACTTTTCCGCAAGCGGAGTGGAAACCTTTAAACAAGTGAGATGS 922
Db 1037 GTGCTTCTGTATAGGCTATTTGCAAGTGGGAAATGGCGTACTTTGCAATGGAAATGG 1096
Qy 923 AAAAATAACAACTGGGAAAGTCCCATACGTTATGACTTTTGGCTCGATAGCTATTTCTCT 982
Db 1097 AGGGTTTTCAGAAAGGATATGTTGCTTATGTTATGACTTTGGTTGGACTTCAATAGCCT 1156
Qy 983 GGCAGTCTACCAATTTGGGCTGCTGGGACTGATCTTTGAGTCAATCTCTGTGTTCTCCA 1042
Db 1157 GGCAGGTATGCTCTGTGTTGGTGTGTTGGCTTGATCTTCCCTAGTGTCTTCTCTACTCCA 1216
Qy 1043 ATTCCATACTGCTGTTGGGATTCGCTATAGTTCAGTCTGTAGCAGTATGTTTCCATG 1102
Db 1217 ATGTTATAGCAAGTTCTTTAGCCGTAACCTCTATGCTGTGTTATAGTATTTTCATG 1276
Qy 1103 ATAAATGAACGCGTCAAAAGATCTTCTCCATCATTTTATGCTATCTGGGATTCATTTCA 1162
Db 1277 ATAAAGATGAATGGGTGAAGATAATTTCTATGCTTTTGGCTCTATGGGTTTGGCTCTT 1336
Qy 1163 TTGCTATCAGCACTACCTCGAGCAAAAGAG 1194
Db 1337 ATATTTATCAGAATTATCTTGATGATTCAAAG 1368

RESULT 6

US-10-437-963-82061/c
; Sequence 82061, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 82061
; LENGTH: 1513
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81525C.1
US-10-437-963-82061

Query Match 22.8%; Score 295; DB 17; Length 1513;
Best Local Similarity 56.9%; Pred. No. 1.8e-76;
Matches 563; Conservative 0; Mismatches 420; Indels 6; Gaps 1;
Qy 211 AAGAACTATAAGAAATGGCTTCGTATTTCCATTTACGTGTTCTTTGCTTCCTGTTGCCAA 270
Db 1270 AAGAGTTGCAATGGTGGCTGATGGTGGGAGTTAAACATGTTCTTCTCATTTGCCGTGAG 1211
Qy 271 GCATTTCTACAATTTTGGGAGAGTTTACTATGAAAATGGTGGGAGAGTACATGGATG 330
Db 1210 ACCGCATCTACATTTCTTGGGAGGTTCTACTACAATCAAGGGGCAATAGCAAGTGGATG 1151

Qy 331 GGAAACACTTGTCCAACTAATCGGCTTCCCTGTTCTGTTTCTCTTCGCTTCTTTTCCCAA 390
Db 1150 TCCACATCGTCCAACTGCTGCTTCCGATTTTGTTCATTTGCCCTATTTCTTTTCCAT 1091
Qy 391 ACCAAATCCCAACCAACAGAGATTTTCAGAAAGTTCTC-----TTCCTTCACC 444
Db 1090 TCAAAGACATCTTCTACAAACAGCTACTAGTAGTCTCTGCCCTTACAAATTTCTATCCCC 1031
Qy 445 ATTCTTGGATCAGTTTACATCGTTACTGACTATTAGTCTGCTAACTCTTATATGTCC 504
Db 1030 AAAATTAATCTGATATATGTTGCTTGGGCTCATCATTTGCTGACAGCACTTGATGAT 971
Qy 505 TCTGTGTTTACTATATCTTACAGTTTCTACTTTCTCCCTCATCTTGGCCTCACAAATG 564
Db 970 TCTATGGCTACTATATCTTCCGGTCTCAACATATTCGCTCATCTGTGCTAGTCAGCTT 911
Qy 565 GCTTCACTGCTTCTTCTCATATTTTCTAACTGCGAAGATTCACACCTTTTCATG 624
Db 910 GCTTCAATGCTGCTTCTCATATTTCTCAATGCTCAAAAATTCACCCCTCTGATTTTC 851
Qy 625 AATTCTGTTTCTCTTACTATTTTCTCTGCTGCTGCTGCTCAACACTGATTCGGA 684
Db 850 AATTCCGTAAGTCTCTTACGTTTCTGCTTCACTCTTGGAGTTGATGAAGATTTCTCAG 791
Qy 685 AACACAGCAAAAGTGTCTAGAGTAAATATGTAGAGGATATATGTACCATTTGGTGT 744
Db 790 GGAATCTACTAGTATATCATATGGAAGTACATTTTGGGTTTCTGTTGACACTAGGGCA 731
Qy 745 TCTGTGGGATGGATGCTGCTATCCCTGTTACAACTGATCCTCAGGAAGTTTAAAG 804
Db 730 TCAGCTCATACTCGCTCATTTCTCTCCCTGATGCAAGTCACATTTGAGAAGGTTATTAAG 671
Qy 805 AAGCAACATTTCTCAACGCTCACTGCTGCTTACCAATCTCTAGTTCGAACTGT 864
Db 670 AGGAGACCTTCTCAGTTGTGTTGAACATGACAGATATATACAGCTCTCTGTCGCAACAT 611
Qy 865 GTGGTTCTCATAGGACTTTTCCAAAGCGGAGTGGAAACCTTTAACTAGTGGAGAA 924
Db 610 GCTTCTCTTGTGGGTATTTGCAAGTGGTGAATGATGACTTTACAGAGAGATGCAT 551
Qy 925 AACTACAACTGGGAAAGTGCATACGTTATGACTTTTGGCTCGATAGCTATTTCTCG 984
Db 550 GCATTTCCAGTCTGGGAAGCTGTATATGTAATGACACTGCTGAGCGCTATATCTTGG 491
Qy 985 CAAGTCTACACCATTTGGGCTGCTGGAGTGTCTTTGAGTCACTCTCTGTTCTTCCAT 1044
Db 490 CAGGTAGCATCAGTTGGAGTGGTGGATGATCTTTTGTGTTGCTCATCGCTGTTTCAAAT 431
Qy 1045 TCCATAACTGCTGTGGGATTTGCTATAGTTCCAGTTGTAGCAGTGAATTTTTCATGAT 1104
Db 430 GTGATAAGCACCTAGCTTACCCATCATCTCTGTTTGTGCTGATTTTCTTTCATGAC 371
Qy 1105 AAAATGAACGCTCAAGATCTTCTCCATCATTTTATGCTATCTGGGATTCATTTCA 1164
Db 370 AAGATGGATGGAGTAAAGATTATAGCTATGCTGATGGCAATTTGGGATTTATGTCATAT 311
Qy 1165 GTCTATCAGCACTACCTCGACGAAAGAA 1193
Db 310 GGCCCAACCAATTATATGTTGATGGCAAGAA 282

RESULT 7

US-10-437-963-43860/c
; Sequence 43860, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 114931
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_74793C.1
US-10-424-599-114931

Query Match 11.8%; Score 152.4; DB 16; Length 522;
Best Local Similarity 57.5%; Pred. No. 2.5e-34;
Matches 293; Conservative 0; Mismatches 216; Indels 1; Gaps 1;

Qy 568 TTCACGCTTCTCTCATATTTTCTAAACGCGAGAGTTTCACACCTTTTCATTTGTGAAT 627
Db 6 TTCATGCGAGTGCTCACTTCTTCTATCAATTTCCCAAAAGTTTCACATGCAATGATCTTAAC 65

Qy 628 TCTCTGTTTCTCTTACTATTTCTCTGCCCTCTCGTGGTCAACACTGATTCGGAAGAAC 687
Db 66 TCTATAATCGTCTTACCATATCATGTTACCTGATTCGACTCAATGCTGATGAGGAA 125

Qy 688 ACAGCAAAAGTCTAGAGTAAATATGTATGATGAGGATAATATGACCATTTGGTGTCTTCT 747
Db 126 ACAAGAACCTTTCCAAACAGAAACAAATAATTTGGGTTCTTCTGTGCCCTAGGTGCTATCT 185

Qy 748 GCTGGATTGGATTGCTCTCATCTCCCTGGTCAACACTGATCCTCAGGAAGTTTAAAGAG 807
Db 186 GCCATATTGATGTCATCACTCTCTTAAGCAATTTTAATTTTGAGAAATATTAAGACT 245

Qy 808 CAACATTTCTCAACGGTCACTGACTTGGTTCGCTTACCAATCTCTAGTTGCAAGCTGTGTG 867
Db 246 GAAACTTTTCTACTGTATTAAGCATGATTTTACCCCAATGATCGTGGGTACAATTGGT 305

Qy 868 GTTCTCATAGACTTTTCGAGCGGGAGTGGAAAACCTTTAAACAGTGAGATGGAAGAAC 927
Db 306 GGCCTTGTGGGATTGCTTTGTAAGTGGAGATTGGAGAACTATGGGAAATGGAGATGAAGGAA 365

Qy 928 TACAACTGGGGAAGTCCCATACGTTATGACTTTGGCTCGATAGCTATTTCTCTGGCAA 987
Db 366 TTCGAAAACGGTAGTGTGTATATGATGATGACTCTGGTTGTACTTCTGTGACATGGGAG 425

Qy 988 GTCTACACCATTTGGCGTGGGACTGATCTTTGAGTCACTCTTCTGTGTTCTTCCAAATTC 1047
Db 426 ATAGTTTGTGTTGATGCTGGGTTGATTTTGAGGTATCTTCAATACCCTCAGCCTGC 485

Qy 1048 A-TAACTGCTGGGATTTGCCATATAGTTCC 1076
Db 486 ACTAAGCAATCTGGAATTTGAACATAGCTCC 515

RESULT 15
US-10-437-963-83900
; Sequence 83900, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 83900
; LENGTH: 2605
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_83188C.1
US-10-437-963-83900

Query Match 10.9%; Score 140.6; DB 17; Length 2605;
Best Local Similarity 48.7%; Pred. No. 2.1e-30;
Matches 460; Conservative 0; Mismatches 464; Indels 21; Gaps 2;

Qy 245 ACGTGTTCCTTTGCTTGTGCCAAGCACTTTCTACAATTTTGGGCAGAGTTTACTATG 304
Db 1331 ACATGTTTCATGGTGCTCTCGGCGCAGACCGTGGCCACCCTCTCGCGCGCTGTACTACA 1390

Qy 305 AAAATGCTGGGAAGATACATGGATGGAAACACTTGTCCAACTAAATCGGCTTCCTCTGTC 364
Db 1391 ACTCCGGCGGCACAGCAAGTGGATGGCAAGCTCACGCACTCGCGCGCTCGCGCTGC 1450

Qy 365 TGTTCCTCTCCGCTTCTTTTCCAAACCAAAATCCAAACCAACAGAGAGATTTCA 424
Db 1451 TCGCATCTCTCTCTCTTCCAGCGCGCGCGCGCGCGAGCCAGCGCGCGCGCG 1510

Qy 425 GAAAGTTCTCTTCTTCAACCATTTCTTGATCAGTTTACATCGTTTACTTGGACTATTAGTGT 484
Db 1511 CAAAGATGGCGCCCATC-----TACGTGGCTCGGATCATCATCG 1552

Qy 485 CTGTAACCTCTATATATCTCTCTGTTTACTATATCTTACCAGTTTCTACTTTCTCTCC 544
Db 1553 GCTTCGACAACCTGATGCTCTGTCAGCGCTGCAGTACCTGCCGTCTCCACCTTCTCGC 1612

Qy 545 TCATCTTTGGCTCACAAATGGCTTCTACTGCTCTTCTCTCATATTTTCTAACTCGCAGA 604
Db 1613 TCGTGGCGGCGACGCACTCGGCTTCAACTCCGCTCACTCCAGGCTCATCAACGCGCAGC 1672

Qy 605 AGTTCAACCTTTCAATTTGTAATTTCTGTTCTCTTCTTCTTACTATTTCTCTGCCCTCTCG 664
Db 1673 GGTTCACGGTGTGATCGCAACTCCGTCGTGTGTCTCACTCTCGCGCGCTGTCTCG 1732

Qy 665 TGGTCAACACTGATTCGGAATAACACAGC---AAAAGTGTCTAGAGTAAATATGTGATAG 721
Db 1733 GCATCGGGGCTCTCTCGACGAGACCGCCAGCAGCGTGCAGGGGCAAGTACCCCGCG 1792

Qy 722 GGATAATATGATCAATTTGGTGTCTCTGCTGGATTTGATTTGCTGCTATTCCTGTGTACAAC 781
Db 1793 GGTTCGCGCTGACGCTCGCGCTCTCCGCTGTTCGCGCTCATCTCTGCTGTGTCGAGG 1852

Qy 782 TGATCTCTCAGAGAGTTTAAAGAACAAACATTTCTCAACGCTCACTGACTTGTGCTGT 841
Db 1853 CCACCTTCGAGAGAGTGGTCCGACCGGACGCTCCGCTGGGTTCTCGCGCGCAGCTGT 1912

Qy 842 ACCAATCTCTAGTTGCAAGCTGTGTGTTTCTCTATAGGACTTTTTCGAAGCGGGAGTGA 901
Db 1913 GGACCAACGCTGGTGGCTCTGACGCTGTCGGGCTGGGCTCTCTCGCTCGGGGACTGA 1972

Qy 902 AAACCTTAAACAGTGAGATGGAATAACTACAACTGGGGGAAGTCCCATACGTTATGACTT 961
Db 1973 GGACGATACCGCGCGAGATGGCGGCTTCAAGGACGGGAGCGGAGGTACGTGGCGACGC 2032

Qy 962 TGGCTCGATAGCTATTTCTTGGCAAGTCTACACCATTTGGCGTGTGGAGTGTCTTTG 1021
Db 2033 TGGTCGGGACGCGGTGTCTGTGGCAGGTGATGGCGGTGGGCTCGCTCGGCTGTATCGTGA 2092

Qy 1022 AGTCATCTTCTGTGTTCTTCCAAATTTCCAACTGCTGTGGGATTCCTATAGTTCAGTTG 1081

Db	2093	GGGTGTCGTCGCTGTTGCGCCAAACGTGACGGGGACGCTGTGCTGCGCGCTGGTGCCGGTGT	2152
Qy	1082	TAGCAGTGAATTGTTTTCATGATAAAATGAACGCGTCAAGAGATCTTCTCCATCATTTTAG	1141
Db	2153	TCGCCGTCGCGCTGTTTCGGGGACAGGATGACCGGGATCAAGGCCGTATCCATGCTCATGG	2212
Qy	1142	CTATCTGGGGATTTCATTTGTTCTATCAGCACTACCTCGACG	1186
Db	2213	CCGTCCTGGGGTTTCCTCTCGTACGCGTACCAGCAGTACATCGACG	2257

Search completed: November 1, 2004, 23:40:02
Job time : 674.694 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:45:29 ; Search time 4606.44 Seconds
(without alignments)
10228.408 Million cell updates/sec

Title: US-09-913-767-4

Perfect score: 1293
Sequence: 1 ggaagctctcttagagtggt.....tatgatcaaacatattccc 1293

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	806	62.3	1340	3	CNS0A29J
2	778	60.2	1431	3	CNS0A2KR
3	767	59.3	1375	3	CNS0A2S8
4	760.4	58.8	1312	3	CNS0A2U8
5	760.4	58.8	1359	3	CNS0A2R3
6	461.2	35.7	699	8	B77316
7	399.2	30.9	691	8	BZ087060
8	361	27.9	709	8	BH498028
9	355.8	27.5	1519	3	CNS0A3B5
10	302	23.4	428	1	AU227210
11	298	23.0	424	5	BP575101
12	290.4	22.5	1323	3	CNS0A36P
13	282.2	21.8	558	1	AV828990
14	273.8	21.2	446	1	A1996743
15	252.2	19.5	782	7	CO105310
16	249.2	19.3	822	8	BZ491815
17	248.4	19.2	922	8	BZ967767
18	245	18.9	498	1	AU226302
19	242	18.7	1374	3	CNS0A0EV
20	241.2	18.7	666	5	BO157452
21	229.4	17.7	770	9	CG344371
22	227.8	17.6	627	5	BQ155273
23	227.2	17.6	837	6	CB292743
24	226.2	17.5	739	7	CO105307

25	223.6	17.3	768	1	AJ795195
26	222.6	17.2	444	1	AV439952
27	218.8	16.9	905	5	BX927581
28	217.6	16.8	630	4	BG525645
29	214.8	16.6	1289	3	CNS0A2HI
30	213.6	16.5	568	4	BMI76948
31	212.2	16.4	668	5	BQ704880
32	210.4	16.3	1299	3	CNS0A41J
33	204.6	15.8	426	5	BP666513
34	202.8	15.7	806	4	BI310188
35	200.8	15.5	706	1	AJ803278
36	200.4	15.5	780	6	CD485871
37	200.2	15.5	701	4	BI204933
38	200.2	15.5	809	7	CF445584
39	199.4	15.4	591	4	BI205616
40	199.4	15.4	742	4	BI205696
41	199.4	15.4	742	4	BI210351
42	199.4	15.4	839	4	BG446722
43	198.4	15.3	726	8	BZ045790
44	198.2	15.3	633	1	A1777352
45	197	15.2	589	7	CK096760

ALIGNMENTS

RESULT 1
CNS0A29J
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

CNS0A29J 1340 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTSL71ZA05 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
BX827536
HTC; GSLT cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1340)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1340)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
Location/Qualifiers
1. 1340
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"

/clone="GSLTSL712A05"	
/tissue_type="Adult vegetative tissue"	
/plasmid="pCMVSPORT_6"	
gene	1. .1340
/gene="At4g18197"	
ORIGIN	
Query Match	62.3%; Score 806; DB 3; Length 1340;
Best Local Similarity	81.9%; Pred. No. 4.6e-215;
Matches 958; Conservative	0; Mismatches 200; Indels 12; Gaps 2;
Qy	124 GATCATACACATAGAGCAAACTTAACAGCTCAGGAGGAAATGAATACACACCGAA 183
Db	65 GGTGACCAAGAACTTAGAAGCAAACTTATAGATCATGAGGTGGTAAC-----GAA 115
Qy	184 ATCGAATCTTCGTCGCTCACTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTCAT 243
Db	116 TCATCATCATCAGCTGTGCCTCAACCGAGAACTATAAAGGTGCTTCGTTCTCCATA 175
Qy	244 TAGTGTTCTTTGCTGCTTGGCAAGCACTTCTACAATTTTGGGAGAGTTTACTAT 303
Db	176 TACGTAATCTTTGTCTCTTTTGGCAGCCACTAGCTACAATTTCTGGGTAGATTGTACTAT 235
Qy	304 GAAATGTTGGGAAGAGTACATGATGGAACACTTGTCCAACTAATCGGCTTCCCTGTT 363
Db	236 GAAATGGAGAAATAGACATATGTGTAACACTTCTTCAACTCATTTGGCTTCCCTGTA 295
Qy	364 CTGTTCTCTTCCGCTCTTTTCCCAACCAAAATCCCAACCAACAGAGATTTTC 423
Db	296 CTGTTCTGTTCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATGTC 355
Qy	424 AGAAGTTCTCTTCTTCCACATTTCTTGGATCAGTTTACATCGTTTACCTGGACTATTAGTG 483
Db	356 AGTCAGTCCCTCTCTTCCACACCTTTCATCGCTTACTTTGTGCACTGGACTGCTAGTG 415
Qy	484 TCTGTAACCTTATATGCTCTGTTGTTTACTATCTTACAGTTTCTACTTTTCTCC 543
Db	416 TCCGCTATGCTTATTTGCTGCAAGTGGTGTCTTACTTACCAGTCTTACTTTTCTCC 475
Qy	544 CTCACTTTGGCTCACAATTTGGCTTCACTGCCTTCTTCTCATATTTTCTAACTCGAG 603
Db	476 CTCACTTTGGCTCACAGTTGGCTTCACTGCCTTCTTCTCATATTTCTTAACTCGAA 535
Qy	604 AAGTTCAACCTTTCAATGTGAATTTCTGTTTCTTCTTACTATTTCTTCTGCTCTCTC 663
Db	536 AAGTTCACTCTTTGATAGTCACTTCTTGTCTTCTCTCACTGTATCTCTGCTCTTCTT 595
Qy	664 GTGGTCAACCTGATTTGGGAAACACAGCAAAAGTGTCTAGAGTAAATATGTATAGGG 723
Db	596 GTGGTCAACCTGATTTCAAGAACTCAACTAATGTATCTAGAGTACATGTATGTATCGG 655
Qy	724 ATAATATGTACATTTGCTTCTGCTGGGATTTGGATTTGCTGCTATCCCTGGTACAACTG 783
Db	656 TTCAATATGATCACTCGGTGCTTCCGCTGGGATTTGACTTACTTATCTCTGATACAAATG 715
Qy	784 ATCTTCAGGAGTTTAAAGAGCAACATTTCTCAAGGTCACGTGCTGCTGCTTAC 843
Db	716 CTCTTCAGGAAGTTTTCACGAACATACATCTCTCAGCAGTCACGGACTTTGGCCATTAC 775
Qy	844 CAATCTCTAGTTGCAAGCTGTGTTCTCATAGACTTTTTCGCAAGCGGGAGTGGAAA 903
Db	776 CAGTCTCTAGTTGGAGTTGTGTAGTTCTCATAGACTTTTTCGAGTGGAGATGGAA 835
Qy	904 ACTTTAAACAGTGAATGAAACACTACAACTGGGAAAGTGCCATACGTTATGACTTTG 963
Db	836 ACTTTGCAAGTGAATGAGAACTACAACTCGGAAAGTGTATATGTTTGTGACTTTA 895
Qy	964 CCCTCGATAGCTATTTCTTGGCAAGTCTACACCAATTTGGCTGCTGGGACTGATCTTTGAG 1023
Db	896 GCCTCGGAGCTATTTCTTGGCAAGTCTACACTCTTGTGCTTGTGGGATTCATCTTCGAG 955
Qy	1024 TCATCTCTGTTCTTCCAAATTCATTAACCTGCTGTGGGATTCCTATAGTTCCAGTTGTA 1083
1015 TCATCTCTGTTCTTCCAAATTCATTAACAGCTGTGGATTGCTATAGTTCCAGTTGCG	
1084 GCAGTGATTTTTCATGATAAAATGAACGCGTCAAAAGTCTTCTTCATCATTTTAGCT	
1016 GCAGTGATAGTTTTCATGATAGATGGACGATCAAAAATCTTCTCCATTTTAGCT	
1144 ATCTGGGATTCATTTTCATTTGCTATCAGACTACCTCGACGAAAGAGTTTGAAGACT	
1076 ATCTGGGCTTCTCTCATTTGCTATCAGACTACCTCGACGAAAGAGTTTGAATACT	
1204 AGCCACAAAGTCTGTAGGAGATCTCATCTACTACTGCTGAGGAGGTCACACAAAC	
1136 AGCCACAAAGTCTGTAGGAGATCTTCAT---CTACCTTTGAGGAGGTCACACAAAC	
1264 ATACATAGTGTATGATCAAAACATATTTCC 1293	
1193 ATACAAAGTGTGTATCAAAAGCATATTTCC 1222	
RESULT 2	
CNS0A2KR	1431 bp mRNA linear HTC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION	GSLTSL692B03 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).
ACCESSION	BX829081
VERSION	BX829081.1 GI:42460228
KEYWORDS	HTC; GSLT cDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE	1 (bases 1 to 1431)
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1431)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPF (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis. Location/Qualifiers
FEATURES	1. .1431
source	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/strain="Col-0"
	/db_xref="taxon:3702"
	/clone="GSLTSL692B03"
	/tissue_type="Silique"
	/plasmid="pCMVSPORT_6"
	1. .1431
	/gene="At4g18205"
gene	
ORIGIN	

Query Match		60.2%;	Score 778;	DB 3;	Length 1431;
Best Local Similarity		78.9%;	Pred. No. 3.6e-207;		
Matches 964;		Conservative 0;	Mismatches 230;	Indels 28;	Gaps 2;
QY	15	AGTGGTCATGGAATAAATCAAGTAATCTATGTCAATGCTAAGCAGATGCATCTCGTAG	74		
DB	39	ATTTCTCATGGGCATATCTCAAGTACACATATTGCAATGTGAAGCAAAAGGTACCTGATGA	98		
QY	75	TGTAGATTACTTGATTTCTTTTCGCTAACTCGT-----TGTTTTTGATGTTTTTCAG	124		
DB	99	CGTTAGATCATTTGATCTCTTTTCGATGAATATAATACCACGAGTGTAAATTTTCAG	158		
QY	125	ATCATTAACAACATAGAAAGCAAACTTAACAGGTGAGGAGGAATGAATACACCATGAAAA	184		
DB	159	GTACACAGAACTTAGAAGCAAACTTCTTAGATCATGAGGAAACT-----	202		
QY	185	TGCAATCTTCGTCGCTGCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTCCATTT	244		
DB	203	--GAATCATTTTCAGTACCTCAACAGAACTGTAAAGAGTGGCTCCGTGTCTCCATAT	260		
QY	245	ACGTGTTCTTTGCTTGTCTTGCCAAAGCACTTTCTACAAATTTTGGCAGAGTTTACTATG	304		
DB	261	ACGCAATCTTTGTCATCTTCTGCCAACCACTTCTGCTACAGTTCTGGGTAGACTGTACTAG	320		
QY	305	AAATGGTGGGAAGAGTACATGATGGGAACAATCTGTCACAACTAATCGGCTTCCCTGTTTC	364		
DB	321	AAATGGAGGGAAGAACATATATGTGTAACACTTCTTCAACTCATTTGGCTTCCCTGTAC	380		
QY	365	TGTTTCTCTTCGCTCTTTTCCCAACCAAAATCCCAAAACCAAGAGCAGATTCA	424		
DB	381	TGATTTCTTTCGCTCTTTTCTCGAATCAGGCAACCCCAAAATCAACAGATACAAATTTCA	440		
QY	425	GAAAGTCTCTTCTTCCATCTTGGATCAGTTTACATCGTTACTGGACTATTAGTGT	484		
DB	441	GTCACTCCCTTCTTCCACCACCTTGATCGTTTACTTGTGCACTGACGTAGTGT	500		
QY	485	CTGTAACTCTTATATGTCCTCTGTTGGTTTACTATACATTAACAGTTTCTACTTCTCC	544		
DB	501	CTGCTTATGCTTATTTGTCGAGTTGGGTGCTTTACTTACCAAGTCTCTACTTCTCC	560		
QY	545	TCATCTTGGCTCACAATTTGGCTTCACTGCTTCTTCTCATATTTCTAACTGCGAGA	604		
DB	561	TCATCTTGGCTCACAAGTTGGCTTCACTGCTTCTTCTCATATTTCTTAACTGCGAAA	620		
QY	605	AGTTCACACTTTTCAATTTGTAATCTCTGTTTCTTCTTACTATTTCTGCTGCTTCTCG	664		
DB	621	AGTTCACCTTTTGTAGTCAATTTCTTGTCTCTTACGTTTCTTCTGCTTCTGCTG	680		
QY	665	TGTCACAACACTGATTCGGAACACAGCAAAAGTGTCTAGAGTAAATATGTATAGGGA	724		
DB	681	TGGTCAACACTGATTCAGAAAACACAAATATGATCTAGAGTACAAATATGTATGGGT	740		
QY	725	TAAATATACATTTGGTCTTCTGCTGGATTTGGATTTGCTATTTCTTCTGCTGCTG	784		
DB	741	TCATCTGTACCAATGGTCTTCCGCTGGATTTGGATTTGCTATCTCTGATACAACTGC	800		
QY	785	TCTCAGGAAGTTTAAAGAAAGCAAACTTCTCAACGGTCACTGACTTGTGCTGCTTACC	844		
DB	801	TCTTCAGGAAGTTTTCAGAAAGCATACATCTCAGCAGTCTGCACTTGGCAATTTACC	860		
QY	845	AATCTCTAGTTCAGCTGTGGTTCTCATAGGACTTTTTCGAAAGCGGGAGTGGAAAA	904		
DB	861	AGTCTCTAGTCTACTTCTGTGTGTACTCATAGGACTGTTTTCGAAGTGGAGATGGAGAA	920		
QY	905	CTTTAACAAGTGCAGATGGAAACTCAAACTGCGGAAGTGCCTACGTTATGACTTTGG	964		
DB	921	TTCTGCGAGTGCAGATGGAAACTCATACTGCGGAAAGTGTCAATATCTTGACTTTGG	980		
QY	965	CCTCGATAGCTATTTCTCGCAAGTCTACACCATTTGGGCTCGTGGGACTGATCTTTGAGT	1024		
DB	981	CCTCAGCAGCTATTTTCTGGCAAGTATACACTGTTGTTGTGGGATTAATCTTCGAGT	1040		
QY	1025	CATCTTCTGTGTTCTCCAATTCATAACTGCTGTGGGATTTGCTATAGTTCCAGTTGTAG	1084		

DB	1041	CTTCTTCTGTGTTCTCCAATTCATAAATGCTGTGGACTGCTATAGTTCCAGTTGTAG	1100		
QY	1085	CAGTCATTTCTTTTCCATGATAAATGAACGGCTGCAAGATCTTCCCATCATTTAGCTA	1144		
DB	1101	CAGTCATGATGTTTTCATGATAAAGATGGATGCAATCAAGATTTTCTCCATCATTTAGCTA	1160		
QY	1145	TCCTGGGATTTCAATTTTCTATCAGCACTACTCTGACGAAAGAGTTGAAGACTA	1204		
DB	1161	TCCTGGGCTTCTTTTCAATTTCTATCAGCACTACTCTGACGAAAGAGTTGAAGACTT	1220		
QY	1205	GCCACACAAGTCTCTGTAGGAGA	1226		
DB	1221	GCCACAAAAACCTGTTGAGGA	1242		
RESULT 3					
CNS0A2S8					
LOCUS					
DEFINITION					
Arabidopsis thaliana Full-length cDNA Complete sequence from clone					
GSLTSL202A10 of Adult vegetative tissue of strain col-0 of					
Arabidopsis thaliana (thale cress).					
ACCESSION					
BX827187.1 GI:42459879					
VERSION					
HTC; GSLT_cDNA.					
KEYWORDS					
Arabidopsis thaliana (thale cress)					
SOURCE					
Arabidopsis thaliana					
ORGANISM					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
Rosids; eurosids II; Brassicaceae; Brassicaceae; Arabidopsis.					
REFERENCE					
1 (bases 1 to 1375)					
AUTHORS					
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,					
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,					
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.					
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:					
A Combined Approach to Evaluate and Improve Arabidopsis Genome					
Annotation					
Unpublished					
2 (bases 1 to 1375)					
Genoscope.					
Direct Submission					
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :					
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr					
- Web : www.genoscope.cns.fr)					
The sequences are based on single pass reads.					
Life Technologies (a division of Invitrogen) members carried out					
full-length libraries construction : Temple G.					
Genoscope members carried out sequencing and annotation : Castelli					
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,					
Schachter V., Weissenbach J., Salanoubat M.					
URGV INRA : Clepet C., Caboche M.					
Annotation is based on the June 2003 version of the Arabidopsis					
genome released by MIPS (Munich information center for Protein					
Sequences). 5 prime and 3 prime are assembled with Phrap.					
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full					
length					
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.					
Location/Qualifiers					
1. .1375					
/organism="Arabidopsis thaliana"					
/mol_type="mRNA"					
/strain="Col-0"					
/db_xref="taxon:3702"					
/clone="GSLTSL202A10"					
/tissue_type="Adult vegetative tissue"					
/plasmid="pCMVSPORT_6"					
1. .1375					
/gene="At4g18205"					
FEATURES					
source					
gene					
ORIGIN					
Query Match 59.3%; Score 767; DB 3; Length 1375;					
Best Local Similarity 83.3%; Pred. No. 4.4e-204;					
Matches 872; Conservative 0; Mismatches 175; Indels 0; Gaps 0;					

Qy	180	GGAAATCGAATCTTCGTCCTGACCTCAATCGAAGAACTATGAAGAAATGGCTTCGTATTTC	239
Db	110	GGAAACTGAATCATTTTCAGTACCTCAACCAAGAACTGAAGAGTGGCTCCGTGCTC	169
Qy	240	CATTACGTGTTCTTTGTCCTTGTGTCGAAGCACTTCTACAAATTTGGGACAGTTTA	299
Db	170	CATATACGAATCTTTGTCACTCTCTGCCAACCACTTGTGTACAGTCTGGGTAGACTGTA	229
Qy	300	CTATGAAATGGTGGGAGAGTACATGATGGGAACACTTGTCCAACTAATTCGGCTTCCC	359
Db	230	CTATGAGATGGGGAAGAACATATGTGTGAACACTTCTTCAACTCATTTGCTTCCC	289
Qy	360	TGTTCTGTTCTCTTCGCTTCTTTTCCAAACCAAAATCCAAACCAACAGAGAGA	419
Db	290	TGTACTGATTTCTTTCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAA	349
Qy	420	TTTCAGAAAGTTCTTCTCTTCCATTCCTGGATCAGTTTACATCGTTACTGACTATT	479
Db	350	TGTGAGTCAGTCCCTTCTTCCACACCTTGCATCGGTTTACTTGTGCATGGACTGCT	409
Qy	480	AGTGTCTGCTAACCTTATATATGCTCTGTGTTGTTTACTATPACTTACAGTTTCTACTTT	539
Db	410	AGTGTCTGCTTATGCTTATTTGTCGAGTTGGTGTGCTTACTTACCAGTCTCTACTTT	469
Qy	540	CTCCCTCATCTGGGCTCAAAATTTGGGCTTCACTGCTTCTTCTCATATTTCTTAAACTC	599
Db	470	CTCCCTCATCTGGGCTCACAGTTGGGCTTCACTGCTTCTTCTCATATTTCTTAAACTC	529
Qy	600	GCAGAAGTTACACTTTCAATGTGAATCTCTGTTCTCTCTTACTTCTCTGCTGCT	659
Db	530	GCAAAGTTCACTCTTTGATGATCAATCTTGTGTTCTCTTACGGTTTCTCTGCTGCT	589
Qy	660	CCTCGTGTCAACACTGATTCGGAACACACAGCAAAAGTGTCTAGAGTAAATATGTGAT	719
Db	590	CCTCGTGTCAACACTGATTCGGAACACACAAATGATCTAGAGTAAATATGTGAT	649
Qy	720	AGGATATATGTACCAATGTGCTTCTGTCGGGATTTGGATGTCTGCTATCCCTGGTACA	779
Db	650	TGGGTTATCTGTATCAATGTGCTTCTGCTGCGGATTTGGATGTGCTATCTCTGATACA	709
Qy	780	ACTGATCTCAGGAAGTTTTAAAGACAGCAAACTTCTCAACGCTCACTGACTTGTGCGC	839
Db	710	ACTGCTCTTCAGGAAGTTTTCAAGAGATACATCTCTCAGAGTCTCTGTACTTGTCCAA	769
Qy	840	TTACCAATCTCTAGTTGCAAGCTGTGTGTTTCTCATAGGACTTTTCGCAAGCGGGAGTG	899
Db	770	TTACAGTCTCTAGTTGCAACTTGTGTGTTACTCATAGGACTGTTTGCAGTGGAGAGTG	829
Qy	900	GAAACCTTTAACTAGATGGAAGAACTACAACTGGGGAAGTGGCATACGTTATGAC	959
Db	830	GAGAACTCTGCCAAGTGAATGAGAACTACAACTGGGGAAGTGTATATATCTTTGAC	889
Qy	960	TTTGGCTCGATAGTATTTCTCGCAAGTCTACACCATTTGGCGTCTGTGGGACTGATCTT	1019
Db	890	TTTGGCTCGACAGCTATTTCTGCAAGTATACACTGTTGTTGTGTGGGATTAATCTT	949
Qy	1020	TGAGTCATCTTCTGTGTTCTCCAAATTCATTAATCTGCTGTGGGATTTGCTATAGTCCAGT	1079
Db	950	CGAGTCTTCTTCTGTGTTCTCCAAATTCATTAATCTGCTGTGGGACTGCTATAGTCCAGT	1009
Qy	1080	TGTAGCAGTATGTTTTCATGATATAAATGAACGGCTCAAGATCTTCTCCATCATTTT	1139
Db	1010	TGTAGCAGTATGATTTTTCATGATATAAAGATGGATGTCATCAAGATTTTCTCATCATTTT	1069
Qy	1140	AGCTATCTGGGATTCATTTCAATTTGTTCTATCAGCACTACCTCGACGAAAGAGTTGAA	1199
Db	1070	AGCTATCTGGGCTTCTCTCATTTTGTCTATCAGCACTACCTCGACGAAAGAGTTGAA	1129
Qy	1200	GACTAGCCACAAAGTCTCTGTAGGAGA	1226
Db	1130	GACTTGCCAGACAAACCTGTTGAGA	1156

RESULT 4	CNS02A2U8	1312 bp	mRNA	linear	HTC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone				
DEFINITION	GSLTLS592F07 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).				
ACCESSION	BX827470				
VERSION	BX827470.1	GI:42460504			
KEYWORDS	HTC; GSLT cDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1 (bases 1 to 1312)				
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.				
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1312)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length				
FEATURES	Location/Qualifiers				
source	1. .1312				
	/organism="Arabidopsis thaliana"				
	/mol_type="mRNA"				
	/strain="Col-0"				
	/db_xref="taxon:3702"				
	/clone="GSLTLS592F07"				
	/tissue_type="Adult vegetative tissue"				
gene	1. .1312				
	/plasmid="pCMVSPORT_6"				
	/gene="At4g18205"				
ORIGIN					
Query Match	58.8%;	Score 760.8;	DB 3;	Length 1312;	
Best local Similarity	83.5%;	Pred. No. 2.4e-202;			
Matches	875;	Conservative 0;	Mismatches 172;	Indels 1;	Gaps 1;
Qy	180	GGAAATCGAATCTTCGTCCTGACCTCAATCGAAGAACTATGAAGAAATGGCTTCGTATTTC	239		
Db	124	GGAACTGAATCATTTTCAGTACCTCAGACGAGAACTGAAGAGTGGCTCCGTGCTC	183		
Qy	240	CATTACGTGTTCTTTGTCTCTTGTGTCGAAGCACTTCTACAAATTTGGGACAGTTTA	299		
Db	184	CATATACGAATCTTTTGTCTCTTCTGCAACCACTTGTGTACAGTCTCGGTAGACTGTA	243		
Qy	300	CTATGAAATGGTGGGAGAGTACATGATGGGAACACTTGTCCAACTAATTCGGCTTCCC	359		
Db	244	CTATGAAATGGAGGGAAGAACATATGTGGTAAACACTTCTTCAACTCATTTGCTTCCC	303		
Qy	360	TGTTCTGTTCTCTTCGCTTCTTTTCCCAAAATCCCAAAATCCCAACAGAGAGAGA	419		

Db 304 TGTAAGTCTCTTCGGCTCTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAA 363
Qy 420 TTTGAGAAAGTCTCTTCCTTCCACCAATCTTGGATCAGTTTACATCGTTACTGGACTATT 479
Db 364 TTTGAGTCAAGTCCCTTCTTCCACCACTTGCATCGGTTTACTTGTGCACTGGACTGCT 423
Qy 480 AGTGTCTGTAAGTCTTATATGTCCTCTGTGTTGTTTATATCTTACCAAGTTTCTACTTTT 539
Db 424 AGTGTCTGCTTATGCTTATTTGTCTGAGTTGGTTGCTTTACTTACCAAGTCTCTACTTT 483
Qy 540 CTCCTCATCTTGGCTCAATTTGGCTTCACTGCTTCTCTCTCATATTTCTTAACTC 599
Db 484 CTCCTCATCTTGGCTCAAGTTGGCTTCACTGCTTCTCTCTCATATTTCTTAACTC 543
Qy 600 GCAGAAAGTTCACACCTTTTCAATTTGTAATCTCTGTTTCTCTTACTTATTTCTCTGCCCT 659
Db 544 GCAGAAAGTTCACCTTTTGTATGATCAATCTTTGTTTCTCTTACGTTTCTCTGCCCT 603
Qy 660 CTTGCTGCTCAACACTGATTCGGAACACAGCAAAAGTGTCTAGAGTAAATATGTA- 718
Db 604 CTTGCTGCTCAACACTGATTCGGAACACACAAATGTATCTAGAGTAAATATGTA 663
Qy 719 TAGGATAATATGTAATGCTTCTGCTGGATTCGATTTGCTGCTTCTCTCTGCTGCTAC 778
Db 664 TGGGTTCTATGTAATGCTTCTGCTGGATTCGATTTGCTGCTTCTCTGCTGCTAC 723
Qy 779 AACTGATCTCAGGAAGTTTTAAAGAAGCAACATTTCTCAAGGTCCTGCTGCTGCTG 838
Db 724 AACTGCTCTCAGGAAGTTCACGAAGCATACATCTCTCAGCAGTCTCTGGACTTGGCCA 783
Qy 839 CTTACCAATCTCTAGTTGCAAGTCTGCTGCTTCTCATAGGACTTTTCGCAAGCGGAGT 898
Db 784 ATTACCAAGTCTCTAGTTGCAAGTCTGCTGCTTCTCATAGGACTTTTCGCAAGCGGAGT 843
Qy 899 GGAAGAACTTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 958
Db 844 GGAAGAACTCTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 903
Qy 959 CTTTGGCTCTGATGATTTTCTGCAAGTCTTACACCAATTCGCTGCTGCTGCTGCTGCT 1018
Db 904 CTTTGGCTCTGATGATTTTCTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 963
Qy 1019 TTGAGTCACTCTGCTGCTTCCAAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078
Db 964 TCGAGTCTCTCTGCTGCTTCCAAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1023
Qy 1079 TTGAGTCACTGATTTTCCATGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1138
Db 1024 TTGAGTCACTGATTTTCCATGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1083
Qy 1139 TAGTATCTGCTGCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1198
Db 1084 TAGTATCTGCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1143
Qy 1199 AGACTAGCCACAAAGTCTCTGAGGAGA 1226
Db 1144 AGACTAGCCACAAAGTCTGAGGAGA 1171

RESULT 5

CNS0A2R3 1359 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTSL52E07 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX827477.1 GI:42459927
VERSION HTC; GSLT_cDNA.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 1359)
AUTHORS Castellani V., Aury J.M., Jallion O., Wincker P., Clepet C.,
Menard M., Cruaud C., Quetier P., Scarpelli C., Schachter V.,
Temple G., Caboche M., Weissenbach J., and Salanoubat M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
JOURNAL Annotation
REFERENCE 2 (bases 1 to 1359)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castellani
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Hanque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
Location/Qualifiers
1..1359
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL52E07"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
1..1359
/gene="At4g18197"
Query Match 58.8%; Score 760.4; DB 3; Length 1359;
Best Local Similarity 79.9%; Pred. No. 3.1e-202;
Matches 937; Conservative 0; Mismatches 221; Indels 14; Gaps 3;
Qy 124 GATCAATCAACATAGAGCAACCTAACAGGTGAGGAAATCAATACCACCATGGA 183
Db 39 GGTGACGAGACTTAGAGCAACCTTATAGATCATGAGGTGGA 89
Qy 184 ATCGAATCTTCGTCCGTACCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTCCATT 243
Db 90 TCATCATCATCAGCTGTCCTCAACCCGAGAACTATAAAGGTGGCTTCGTCTCCATA 149
Qy 244 TACGTGTTCTTTGCTTGTTCCTTGGCAAGCACTTTCACAAATTTGGGAGAGTTTACTAT 303
Db 150 TACGTAATCTTTGCTTCTTTTGGCAGCCACTAGCTACAAATTTCTGGGTAGATTGACTAT 209
Qy 304 GAAATGTTGGAGAGTACATGGATGGAACTTGTCCAACTAATCGGCTTCCTCTGT 363
Db 210 GAAATGAGGAATAGCAATATGTGTATACACTTTTCAACTCATTTGGCTTCCTGTA 269
Qy 364 CTGTTTCTTCCTCGCTTCTTTTCCAAACCAAAATCCAAACCAAGAGAGATTC 423
Db 270 CTGTTTCTTCCTCGCTTCTTTTCCAAACCAAAATCCAAACCAAGAGATTC 329
Qy 424 AGAAGTTCTCTTCCTTCCATTCAGTTCAGTTTACATCGTTTACCTGGACTATTAGTG 483
Db 330 AGTCAGTCCCTTCTTCCACCCCTTCATCGGTTTACTTGTGCACTGGACTGTAGTG 389
Qy 484 TCTGTAATCTTATATGCTCTGTTGTTTACTATATCTTACCAAGTTTCTTCTTCTCC 543
Db 330 TCCGCTTATGCTTATTTGCTGAGTGGTGTCTACTTACCAGTCTTACTTTTCTCC 449
Qy 544 CTCATCTTGGGCTCAAAATGGCTTCTACTGCTTCTTCTCATATTTTCTAACTCGCAG 603

Db 450 CTCACTTTGGCTCACAGTTGGCTTCTCACTGCTTTTCTCATATTTCCCTTAACTCGCA 509
Qy 604 AAGTTCACACCTTTCAATGTAATCTCTGTTTCTCTTACTATATTTCTCTGCGCTCCTC 663
Db 510 AAGTTCACCTCTTGATAGTGAGTCTTTGCTTCTCTCTCACTGATCTCTGCTCTTCTT 569
Qy 664 GTGGTCAACACTGATTCGAAACACACGAAAGTGCTAGAGTAAATATGTCATAGGG 723
Db 570 GTGGTCAACACTGATTCGAAACACTCAATATGATCTAGAGTACAGTATGTCATCGGG 629
Qy 724 ATAATATGTACATTTGCTTCTGCTGGGATTTGGATTGCTATCTCTCTCTGTTCAACTG 783
Db 630 TTCAATAGTACATCGTGCTTCTGCTGGGATTTGGAGTGTACCTATCTCTAGACAAATG 689
Qy 784 ATCTTCAGGAAGTTTAAAGAACCAACATTTCTCAACGGTCACTGACTTGGTCTGCTTAC 843
Db 690 ATCTTCAGGAAGTTTTCACGAAGCATATACCTCAGCAGTAACGGACTTGGCCATTAC 749
Qy 844 CAATCTCTAGTTCGAAGCTGTGGTCTCTCATAGGACTTTTCGCAAGCGGGAGTGGAA 903
Db 750 CAGTCTCTAGTTCGCGGTGTGTAGTTCTCATAGGACTTTTTCGAAGTGTAGAGTGGAA 809
Qy 904 ACTTTAAACAGTGAGATCGAAACACTACAACTGGGAAAGTCCATAGCTTTATGACTTTG 963
Db 810 ACTTTGCCACGTGAGATGAGAACTACAACTCGGAAAGTGTATATGTTTGTACTTAA 869
Qy 964 GCCTCGATAGCTATTTCTGCGCAAGTCTACACCAATTCGCGTGTGGGACTGATCTTT--G 1021
Db 870 GCCTCGGCAACTATTTCTGCAAGTCTACACTCTTGGTCTTGTGGGATATGATCTTCGC 929
Qy 1022 AGTCATCTCTGTTTCTCCAAATCCATTAACCTGCTGTGGGATTTGCTATAGTTCAGTTG 1081
Db 930 AGTCATCTCTGTTTCTCCAAATCCATTCCTCCCGCTGTGGGACTGCTATGCTTCCCGTTG 989
Qy 1082 TAGCAGTCAATTTGTTTCCATGATAAATGAAGCGTCAAGATCTTCTCCATCATTTTAG 1141
Db 990 CGGCGTGATCGTTTCCATGATAGATAGATGAGCATCAAAATCTTCTCCATCATTTTAG 1049
Qy 1142 CTATCTGGGATTCATTTCTATGCTATGAGTCTACCTCGAGCAAAAGTGTGAAGA 1201
Db 1050 CTATCTGGGCTCTCTCTCAATCGTCTATCAGACTACCTCGAGCAAAAGTGTGAATA 1109
Qy 1202 CTAGCCACAGTCTCTGAGAGTCTCTATCTACTACTCTGCTGAGCAAGGTACACAA 1261
Db 1110 CTAGCCACAGTCTGCTAGAGATCTTCAT---CTACCTGTTGAGGAAGTACACAA 1166
Qy 1262 ACATACATAGTATGATCAAAACATATTTCC 1293
Db 1167 ACATACAAAGTGTGATCAAGCATATTTCC 1198

RESULT 6
LOCUS B77316/c 699 bp DNA linear GSS 16-JAN-1998
DEFINITION T3213TF TAMU Arabidopsis thaliana genomic clone T3213, genomic survey sequence.
B77316
VERSION B77316.1 GI:2773955
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 699)
REFERENCE Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K., and Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and Venter,J.C.
AUTHORS A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genomic Sequencing. Update 3 Unpublished (1997)
TITLE Arabidopsis Genomic Sequencing. Update 3
JOURNAL Unpublished (1997)
COMMENT Contact: Steve Rounsley

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13-21
Class: BAC ends
High quality sequence stop: 699.
FEATURES
Location/Qualifiers
source 1..699
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="T3213"
/sex="hermaphrodite"
/clone_lib="TAMU"
/notes="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII; Produced by Rod Wing"

ORIGIN
Query Match 35.7%; Score 461.2; DB 8; Length 699;
Best Local Similarity 85.4%; Pred. No. 3.7e-118;
Matches 514; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
Qy 625 AATTCTCTGTTTCTCTTACTATTTCTCTGCGCTCTCTGTCACACACTGATTCGAA 684
Db 698 AATTCTTTGTATCTCTTACGGTTTCTTTGCGCTCTCTGTCACACACTGATTCGAA 639
Qy 685 AACACACAAAGTCTAGAGTAAATATGTATAGGATAATATGATACCATTTGGTGTCT 744
Db 638 AACACAAATGATCTAGAGTAAATATGTATAGGTTTCTCTGTCACATTTGGTGTCT 579
Qy 745 TCTGCTGGGATTTGGATTTGCTATCTCTGTCACACTGATCTCTCAGAAAGTTTAAAG 804
Db 578 TCGCTGGGATTTGGATTTGCTATCTCTGATACACTGCTCTTCAGAAAGTTTTCACG 519
Qy 805 AAGCAAACTCTCAACGGTCACTGACTTGGTGTCTTACCAATCTCTAGTTGCAAGTGT 864
Db 518 AAGCATACATCTTCAGCAGTCTCGGACTTGGCCAAATTACCACTCTCTAGTTGCAACTGT 459
Qy 865 GTGGTCTCTCATAGGACTTTTCGCAAGCGGGAGTGGAAACTTTTAAACAAGTGAGATCGAA 924
Db 458 GTGGTACTCATAGGACTTTTCGCAAGTGAGAGTGGAGAACTCTGCCAAGTGAGATGAGA 399
Qy 925 AACTACAACTGGGAAAGTCCCATACGTTATGACTTTGGCCTCGATAGCTATTTCTCTGG 984
Db 398 AACTACAACTGGGAAAGTGTATATATCTTGACTTTGGCCTCAGCAGCTATTTCTCTGG 339
Qy 985 CAAGTCTACACCATTTGGCGTCTGGGACTGATCTTTGAGTCACTCTCTGTTGTTCTCCAAT 1044
Db 338 CAAGTATACACTGTGTGGTGTGGGATTAATCTTCGAGTCTTCTCTGTGTCTCCAAT 279
Qy 1045 TCCATAACTGCTGGGATTTGCCATATAGTTCCAGTTCTAGCAGTATTTGTTTTCATGAT 1104
Db 278 TCCATAACTGCTGGGACTGCCCTATAGTTCCAGTTCTAGCAGTATTTGTTTTCATGAT 219
Qy 1105 AAAATGAACCGGTCAAAAGATCTTCTCCATCATTTTACGATCTCTGGGATTCATTTCAATT 1164
Db 218 AAGATGATGATCAAAAGATTTTCTCCATCATTTTACGATCTCTGGGCTTCTTCTCATTT 159
Qy 1165 GTCTATCAGCACTACCTCGAGGAAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1224
Db 158 GTCTATCAGCACTACCTCGAGGAAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 99
Qy 1225 GA 1226
Db 98 GA 97
RESULT 7
BZ087060

LOCUS
DEFINITION BZ087060 691 bp DNA linear GSS 10-OCT-2002
1kh10d12.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
sequence.
ACCESSION BZ087060
VERSION BZ087060.1 GI:23719662
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 691)
Delehaency,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
Nash,W., Rabinowicz,P.D. and Wilson,R.K.
TITLE Whole genome shotgun reads from Brassica oleracea
JOURNAL Unpublished (2002)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: 1kh10 row: d column: 12
Seq primer: -21UpPOT forward
Class: shotgun
High quality sequence start: 18
High quality sequence stop: 491.
FEATURES
Location/Qualifiers
1..691
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.Oleracea002"
/notes="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea TO1000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."
ORIGIN
Query Match 30.9%; Score 399.2; DB 8; Length 691;
Best Local Similarity 77.4%; Pred. No. 1e-100;
Matches 497; Conservative 0; Mismatches 143; Indels 2; Gaps 1;
QY 492 CTCCTATATGTCCTCTGTGGTTTACTATACCTACCTACCTCTCTCCCTCATCTT 551
DB 36 CGCTTATCTATGCAATGGTTTACTCTACCTACCTGTTCCACCTCTCCCTCATCTT 95
QY 552 GGCCTCACAAATGGCCTTCACCTGCCTTCTCTCATATTTCTAAACTCGCAGAAGTTAC 611
DB 96 GGCCTCACAGTAGCCTTCACCTGCCTTCTCTCTACTTCTTAACTCGCAAAAGTTAC 155
QY 612 ACCTTTCATGTGAAATCTCTGTCTTCTCTTACTACTATTTCTGCTGCTCTCTGCTCAA 671
DB 156 TCCCTTCATAGTCAACTCTCTGTCTCTCTTACCGTCTCTCTGCTGCTCTCTGCTGGA 215
QY 672 CACTGATTCGGAACACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGATAATATG 731
DB 216 CACCGAGTCACAGACACACAAAGTCTCTAGACTAGATGATGATCGGTTTATATG 275
QY 732 TACCAATGTGCTTCTGCTGGGAATGGATGCTGCTATCCCTGGTACAACTGATCCTCAG 791
DB 276 CACCATCGGTGTTACGTGGGATAGGACTGCTGCTATCTCTGATACAACTCCTCTTCAG 335
QY 792 GAAGGTTTTAAGAACCAACATTTCAAGGTCAGTACTGTTGGTGGCTTACCAATCTCT 851
DB 336 GAAGGTTTTCAAGGAGACATACATCGTCACTAGTCTAGGACTTGAATCTACCACTCTCT 395
QY 852 AGTTGCAAGCTGTGTGTTTCTCATAGGACTTTTTCGCAAGCGGGAGTGGAATTTAAC 911
DB 396 AGTTGCAAGCTGTGTGTTTCTCATTTGAGCTTTTTCGTGACGGGAGAGTGGAATTTGCC 455
QY 912 AAGTGAGATGGAATACTACAAACTGGGAAAGTGCCATACGTTATGACTTTGGCCTCGAT 971

DB 456 TAGTGAGATGAGAAACTACAAACTTGGGCAAGTGTCTTATGTTGACCGTGTGCTCTCTGC 515
QY 972 AGCTATTTCTCGCAAGTCTACACCATTTGGCGTCTGGGACGATGCTTTGAGTCATCTTC 1031
DB 516 AGCTATCTCTTGGCAAGTCTACACGTGGTGGCCCTTGGGGGGTTGAAATTTGAAATCGTCTTC 575
QY 1032 TGTGTTTCTCCAAATTCATAACTGCTGTGGGATTGCTATAGTTCCAG--TTGTAGCAGTG 1089
DB 576 TGGGGTCTCCAACTCCATAACGCCGTGGGGTGCCTATAATAACGGGCCATAAGCTGGA 635
QY 1090 ATTGTTTTCATGATAAATGAACCGCTCAAGATCTTCTCTCC 1131
DB 636 TAAATTTTCCATGAAAAAATGAAGCGTCCAAAAATCTTCTTC 677
RESULT 8
BH498028 709 bp DNA linear GSS 13-DEC-2001
LOCUS BOHQK68TF BOHQ Brassica oleracea genomic clone BOHQK68, genomic
survey sequence.
ACCESSION BH498028
VERSION BH498028.1 GI:17706132
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 709)
Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOHQK68TF
Contact: Chris Town
TIGR
7712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: 1P
Class: sheared ends.
FEATURES
Location/Qualifiers
1..709
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone_lib="BOHQ"
/note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"
ORIGIN
Query Match 27.9%; Score 361; DB 8; Length 709;
Best Local Similarity 76.7%; Pred. No. 5.8e-90;
Matches 455; Conservative 0; Mismatches 135; Indels 3; Gaps 1;
QY 189 ATCTTGTCCGTACCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTCCATTACGT 248
DB 117 ATCATCATCAGTACCTCAACCAAGAACTATAAAGATGGATCCGCTCTCTCCATCTATGT 176
QY 249 GTTCTTTGCTCTGCTTGCACGACCTTTCTACAAATTTGGCGAGAGTTTACTATGAANA 308
DB 177 CTCTCTGCTCTCTTCTTCCAAACCACTTGCACAAATCTCGGTAGACTGTACTACGAAA 236
QY 309 TGGTGGGAAGAGTACATCGATGGGAACACTTGTCCAATAATCGGCTTCCCTGTTCTGTT 368
DB 237 CGGTGGAGAGACATACGTGGTAAACGCTTCTCCACTCATTTGGCTTCCCTGTTTCAT 296
QY 369 TCTCTTCGCTTCTTTTCCCAACCA---AAAAATCCCAACCAACAGAGCATTTTCAG 425


```
Db 297 CCTCTTCAACTTCTTTTTCACAACTCAGACAAACAAACAAATCAACAGATACAAATGTCAA 356
Qy 426 AAAGTTCTTCTTCCCTTCCACATCTTGGATCAGTTTACATCGTTACTGGACTATTAGTGTG 485
Db 357 CCAGTCCCTTCTTCTTCAACCCCTTTTATTAGTTTACATGTGCGACCGACTTACTCTC 416
Qy 486 TGCTAACTCTTATATGTCCTCTGTGTGGTTTACTATACCTTACCAGTTTCTACTTCTCCCT 545
Db 417 TGTTCACGCTTATCTCTATGCAATGGTTTACTCTTACCTACTGTTTCCACCTTCTCCCT 476
Qy 546 CATCTTGGCCCTCACAAATGGCCCTTCACTGCCTTCTTCTCATATTTCTTAAACTCGCAGAA 605
Db 477 CATCTTGGCCCTCAGAGTACGCTTCACTGCCTTCTTCTTACTTCTTAACTCGCAAA 536
Qy 606 GTTCACACTTTCATGTGAATCTCTGTTTCTCTTCTTACTATTTCTCTGCGCTCCCTCGT 665
Db 537 GTTCACTCCCTTCATAGTCAACTCTCTGTTTCTTCTTACCGTCTCTCTGCGCTCCCTCGT 596
Qy 666 GGTCAACTGATTCGGGAACACAGCAAAAGTGTCTAGAGTAAATATGATAGGAT 725
Db 597 GGTGACACCGAGTCAAGAACAACAATGTCTCTAGACTACAGTATGTATCGGGGT 656
Qy 726 AATATGTACCATTTGCTGCTTCTGCTGGATGGATTGCTGCTATCCCTCGGTAC 778
Db 657 TATATGCACCATCCGTGCTTCACTGGATAGGACTGCTGCTATCTCTGATAC 709

CNS0A3B5 1519 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSJTF8632D12 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress) .
BX826795
HTC; GSLT cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1519)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1519)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life technologies (a division of invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich information center for Protein
Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_projet_EP/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
Location/Qualifiers
1..1519
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"

FEATURES
source
```

```
/db xref="taxon:3702"
/clone="GSLTF8632D12"
/tissue_type="Flowers and buds"
/plasmid="pCMVSPORT_6"
1..1519
/gene="At4g18210"

Query Match 27.5%; Score 355.8; DB 3; Length 1519;
Best Local Similarity 61.5%; Pred. No. 2.le-88;
Matches 691; Conservative 0; Mismatches 417; Indels 15; Gaps 7;

Qy 136 ATAGAACAACACCTTAAACAGGTTCAGGAGGAAATGAATACCACCACCATGGAAATCGAATCTCG 195
Db 123 AAAGAACAATTCACAGATTCAGATGAAGAATAATTCAGTTAGTAGCAGCCAA--GCA 179
Qy 196 TCCGTACTCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTTCCATTTTA-CGTGTTCTT 254
Db 180 GAAGTATCTTACTTAACACATACAAACGGCTCAGGCGTGACTTTATATACATTTCTT 239
Qy 255 TGTCTTGTCTGGCAAGCACTTTCTACAATTTTGGGCAGAGTTTACTATGAAATGTGG 314
Db 240 TGTCAATTCAGGGCAACACAGTTGCTACAATTTTGGGCAGAGTATACTATGACAAACGGAG 299
Qy 315 GAAGAGTACATGG-ATCGGAACACTTCTCCAACCTAAATCGGCTTCCCTGTTTCTGTTCTCT 373
Db 300 AAACAGTAATATGGCTTAGCAACGGTAGTTCAACTTGTGGCTTCTTCTGTTTCTACTTCAT 359
Qy 374 TCCGCTTTTCCCAACCAAAATCCCAACCAACAGAAAGCAGATTTTCAGAAAGTTCT 433
Db 360 ATTATATCTTGTTCATTTTAAACACATGCAACAACCTGATAGAGATGGAAGAAACCTCAC 419
Qy 434 CTTCCTTACCATCTCTGGATCAGTTTACATCGTTACTGCACTATTAAGTGTCTGCTAACT 493
Db 420 CT-----AGGAACCGGTATTCGTTACGTAGTGTGGACTTCTTGTAGGACAGATT 473
Qy 494 CTATATGTCCTCTGTTGGTTTACTATACCTACAGTTTCTACTTCTCCCTCATCTTGG 553
Db 474 GCTATTTGTACTCCATTTGGACTTCTTACTTACCGTTTCTACTATTCCTGATCTGTG 533
Qy 554 CCTCACAAATTTGGCTTCACATGCTTCTTCTCATATTTTCTTAAACTTCGCAAGTTCAAC 613
Db 534 CATCTAGTTAGCTTCAATGCTTTTTTCTTATTATTTCTTAACTCACAAAACTTACCC 593
Qy 614 CTTCATTGTGAATCTCTGTTTCTCTTACTATTTCTCTGCGCTCCTGCTGCTCAACA 673
Db 594 CTATCATTTTAAATCTCTTTTCTCTCTTAACTATATTTTCCACCTTACTTGTCAATCAATA 653
Qy 674 CTGATTCGAAAAACACAGCAAAAGTGTCTAGAGTAAATATGTATAGGGATATATATGTA 733
Db 654 ATGAGGAGACAGATTCCACAAAGTTACAAAGAGAGATTTTTTAAAGGTTTCATATGCA 713
Qy 734 CCATTTGCTGCTTCTGCTGGGATTTGATTTGCTGCTATCCCTGGTACAACATGATCCTCAGGA 793
Db 714 CCGTTGCTGCTTTTGTGTTTATGTTTATTTTATCCCTACACAGCGAGCCCTTTTATA 773
Qy 794 AG-GTTTTAAAGAAGCAAAATCTCTCAACGCTACTGACTTTGGTCCGCTTACCATCTCTA 852
Db 774 AAAGTCTTAAAGAAGCAAAATTTCTTCAGAAAGTTATGGATATGATAATTAATCAATGAATCTG 833
Qy 853 GTTCAAGCTGTGTGGTTCTCATAGGACTTTTCCGAAGCGGGAGTGGAAACCTTTAAACA 912
Db 834 ATGCGAGTTGTGTTCCGCTGGTGGGCGCTTTTGTCTACAATGATGAAATTTGAAC 893
Qy 913 AGTGAGATGGAATACTACAACTGGGAAAAGTGCCATACGTTATGATCTTTGGCCCTCGATA 972
Db 894 AATGAAAGGAAAACATAAAAAAAGGGAAGTAACTTACAATATGAACCTTAATGTGAACA 953
Qy 973 GCTATTTTCTGGCAAG--TCTACACCATTTGGCGTGGGGAGCTGATCTTTGAGTCATCTT 1030
Db 954 GCTGTTACCTGGGCCGCTATCTCCATCGGTGGAACAAGAACTGATCTTTCAGGCTCTCT 1013
Qy 1031 CTGTGTTCTCCAATTCAT-AACTGCTGTGGGATTCCTATAGTTTCCAGTTGTAGCAGTG 1089
```



```

Db      1014  CTTCTATTCTAAATGCAATAAAGCGTTTGGAACTCCCAAGTGGTTCCTATCTCTGGCTGTA 1073
QY      1090  ATTGTTTTCATGATATAAATGAACGGCTCAAGAGATCTTCTCCATCATTTTAGCTATCTCTGG 1149
Db      1074  ATCATTTTCCATGACAAATAAATAGGTTAAAGTGATTTCTATGATCTCTAGCTATTGG 1133
QY      1150  GGATTCATTTCATTTGCTATCAGCACTACCTCGACGAAAGAGTTGAAGACTACCCAC 1209
Db      1134  GGTTTAACTTCTCTATGCTACCAAAAATATCTTGATGACAAAACTTGAAGAAAATCAT 1193
QY      1210  ACAAGTCCTGTAGGAGATCTCTATCTTACTTACTACCTGCTGAGGAG 1252
Db      1194  GAAATAACAAACAGAAATCCCTGACCCACCAGAAAGCAGAAG 1236

```

```

RESULT 10
LOCUS   AU227210/c
DEFINITION AU227210 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-88-B12 3',
mRNA sequence.
ACCESSION AU227210
VERSION   AU227210.1 GI:19741857
KEYWORDS EST
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 428)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A.,
Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekic@tc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further
details.

```

```

FEATURES
source
1..428
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL14-88-B12"
/tissue_type="root"
/lab_host="DH10B"
/clone_lib="RAFL14"
/note="Site_1: BamHI; Site_2: SalI"

```

```

ORIGIN
Query Match 23.4%; Score 302; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 2e-73;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      992  ACACCAATTCGGTCGCGGACATCTTTAGTCATCTTCTGTGTTCTCCAATTCATAA 1051
Db      428  ACACCAATTCGGTCGCGGACATCTTTAGTCATCTTCTGTGTTCTCCAATTCATAA 369

QY      1052  CTGCTGTGGATTGCCATAGTTCAGTTGTAGCAGTATGTTTCCATGATAAATGA 1111
Db      368  CTGCTGTGGATTGCCATAGTTCAGTTGTAGCAGTATGTTTCCATGATAAATGA 309

```

```

QY      1112  ACGCGTCAAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCATTTTCATTGTCTATC 1171
Db      308  ACGCGTCAAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCATTTTCATTGTCTATC 249
QY      1172  AGCACTACCTCGACAAAAGAGTTGAAGACTAGGCACACAAGTCTCTGTAGAGATCCCTC 1231
Db      248  AGCACTACCTCGACAAAAGAGTTGAAGACTAGGCACACAAGTCTCTGTAGAGATCCCTC 189
QY      1232  ATCTACTACCTCTCAGGAAGGTACAAAACATACATAGTGTATGATCAAAACATATTT 1291
Db      188  ATCTACTACCTCTCAGGAAGGTACAAAACATACATAGTGTATGATCAAAACATATTT 129
QY      1292  CC 1293
Db      128  CC 127

```

```

RESULT 11
LOCUS   BP575101/c
DEFINITION BP575101 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-90-L09 3',
mRNA sequence.
ACCESSION BP575101
VERSION   BP575101.1 GI:48990867
KEYWORDS EST
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 424)
Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,
Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,
Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
Functional annotation of a full-length Arabidopsis cDNA collection
Science 296 (5565), 141-145 (2002)
21932900
11910074
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekic@tc.riken.go.jp
reversed clone; Please visit our web site
(http://pfweb.gsc.riken.go.jp/) for further details.

```

```

FEATURES
source
1..424
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL14-90-L09"
/tissue_type="root"
/lab_host="DH10B"
/clone_lib="RAFL14"
/note="Site_1: BamHI; Site_2: SalI"

```

```

ORIGIN
Query Match 23.0%; Score 298; DB 5; Length 424;
Best Local Similarity 100.0%; Pred. No. 2.6e-72;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      996  CATTGGCGTCGCGGACTGATCTTTGAGTCATCTTCTGTGTTCTCCAATTCATACTGC 1055
Db      424  CATTGGCGTCGCGGACTGATCTTTGAGTCATCTTCTGTGTTCTCCAATTCATACTGC 365

QY      1056  TGTGGATTGCTATAGTTCAGTTGTAGCAGTATGTTTCCATGATAAATGAACGC 1115
Db      364  TGTGGATTGCTATAGTTCAGTTGTAGCAGTATGTTTCCATGATAAATGAACGC 305

```

Qy 1116 GTCAAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCATTTCAATTTGCTATCAGCA 1175
|||||
Db
304 GTCAAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCATTTCAATTTGCTATCAGCA 245
|||||
Qy 1176 CTACCTCGACGAAAGAAAGTTGAAGACTAGCCACACAAAGTCTCTGTAGGAGATCCTCATCT 1235
|||||
Db 244 CTACCTCGACGAAAGAAAGTTGAAGACTAGCCACACAAAGTCTCTGTAGGAGATCCTCATCT 185
|||||
Qy 1236 ACTACTCTCTGAGGAGGTGACACAAACATACATAGTGTATGATCAAAACATATTTCC 1293
|||||
Db 184 ACTACTCTCTGAGGAAGGTGACACAAACATACATAGTGTATGATCAAAACATATTTCC 127
|||||

RESULT 12
CNS0A36P 1323 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSTJFB642P05 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
ACCESSION BX826811 GI:42460388
VERSION BX826811.1
KEYWORDS HTc; GSLT_cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
1 (bases 1 to 1323)
Castelli,V., Aury,J.M., Jallou,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quekier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1323)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
Location/Qualifiers
1..1323
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strains="Col-0"
/db_xref="taxon:3702"
/clone="GSLTJFB642P05"
/tissue_type="Flowers and buds"
/plasmid="pCMVSPORT_6"
1..1323
/gene="At4g18190"

gene
ORIGIN
Query Match 22.5%; Score 290.4; DB 3; Length 1323;
Best Local Similarity 55.7%; Pred. No. 5.1e-70;
Matches 583; Conservative 0; Mismatches 451; Indels 12; Gaps 1;
Qy 181 GAAATCGAATCTTCTCGTCCGATCCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTCC 240
|||||

LOCUS AV828990 558 bp mRNA linear EST 01-APR-2002
DEFINITION AV828990 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-36-N15 5',
mRNA sequence.
ACCESSION AV828990
VERSION AV828990.1 GI:19871050
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 558)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Ono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9080
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.
FEATURES
source Location/Qualifiers
1..558
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL09-36-N15"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAFL9"
/note="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"
ORIGIN
Query Match 21.8%; Score 282.2; DB 1; Length 558;
Best Local Similarity 79.4%; Pred. No. 7.9e-68;
Matches 332; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 180 GGAAATCGAATCTTCGTCGTAACCTCAATCGAAGACTATAGAAATGCTTCGTATTTTC 239
DB 141 GGAATCGAATCATTTTCAGTACTCAACGAAAGAACTGTAGAGTGGCTCCGTTCTC 200
QY 240 CATTTACGTTGCTTTGTGCTTCCTGCTGCAAGCACTTCTACAAATTTGGGAGAGTTTA 299
DB 201 CATATACGCAATCTTTGTCTCTCTGCGCAACCACTTGCTACAGTTCGGTAGACTGA 260
QY 300 CTATGAAATGGTGGGAAGATACATGGATGGGAACATCTTGCCAACTAATCGGTTCCC 359
DB 261 CTATGAAATGGAGGGAAGACATATGTGGTAAACACTTCTTCAACTCATTTGGGTTCCC 320
QY 360 TGTTCGTGTTCTCTTCCTTCCTTTTCCCAACCAAAATCCCAACCAACAGAGAGAGA 419
DB 321 TGTACTGATCTCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 380
QY 420 TTTTCAGAAAGTTCTCTCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 479
DB 381 TTTTCAGTCACTCCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 440
QY 480 AGTGTCTGTAACCTTATATATGTCCTCTCTGTTGTTTATATATCTTACCAAGTTTCTACTTT 539

DB 441 AGTGCTGCTGTTATGCTTATTTGCTGCGACTGGCTTGCTTTACTTACCANTCTCTACTTT 500
QY 540 CTCCTCTCATCTGGGCTCACAATTCGCTTCACTGCTTCTTCTCTATATTTCTTAAC 597
DB 501 CTNCTCATCTCTGGGCTCACAGTTGGCTTCACTGCTTCTTCTCATATTTCTTAC 558
RESULT 14
AI996743/c
LOCUS AI996743 446 bp mRNA linear EST 08-SEP-1999
DEFINITION 701668184 A. thaliana, Columbia Col-0, root-1 Arabidopsis thaliana
cDNA clone 701668184, mRNA sequence.
ACCESSION AI996743
VERSION AI996743.1 GI:5843648
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 446)
AUTHORS Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,
Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Brzoska,P.,
Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D.,
Tan,R., Rose,M., Warren,B., Ton,B., Kastury,K., Borillo,C.,
Carpio,T., Policky,J., Suzuki,G., Argentine,C., Shah,S.,
Nobrega,A., Murry,L., Turner,C., Krikorian,S., Elder,L. and
Hanson,D.
TITLE Arabidopsis thaliana Gene Expression MicroArray
JOURNAL Unpublished (1999)
COMMENT Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.
FEATURES
source Location/Qualifiers
1..446
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="701668184"
/tissue_type="root"
/dev_stage="4 - 7 weeks"
/clone_lib="A. thaliana, Columbia Col-0, root-1"
/note="Vector: pSPORT; Site 1: NotI; Site 2: SalI; cDNA
library was derived from untreated root tissue from
Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks.
Plants were grown in 1:1:1 peat moss/vermiculite/perlite
soil at 22 deg. C +/- 3 deg. C under constant light, and
watered with fertilizer. cDNA synthesis was initiated
using a NotI-oligo(dT) primer. Double-stranded cDNA was
blunted, ligated to SalI adaptors, digested with NotI,
size-selected, and cloned into the NotI and SalI sites of
the pSPORT vector."
ORIGIN
Query Match 21.2%; Score 273.8; DB 1; Length 446;
Best Local Similarity 89.5%; Pred. No. 1.7e-65;
Matches 306; Conservative 0; Mismatches 33; Indels 3; Gaps 1;
QY 952 GTTATGACTTTGGGCTCGATAGCTATTTCTCGGCAAGTCTACACATTTGGCGTCTGGGA 1011
DB 446 GTTTTGACTTTAGCCTTCGAGCTATTTCTCGGCAAGTCTACACTCTTTGGTCTTGTGGGA 387
QY 1012 CTGATCTTTGAGTCACTCTTCTGTGTTCTCCCAATTCATTAATCGCTGTGGGATTGCTATA 1071
DB 386 TTGATCTTCGAGTCATCTCTGTGTTCTTCCCAATTCATTAACAGCTGTGGGATTGCTATA 327
QY 1072 GTTCAGTTGTAGCAGTGAATTTTTCATGATAAATGAACGGCTCAAAAGATCTTCTCC 1131

Db 326 GTTCAGTTGGCGAGTCATAGTATTTTCATGATAGATGGACGATCAAAAATCTTCTCC 267
 Qy 1132 ATCATTTTAGCTATCTGGGATTCATTTTCATTTGTCATCAGCACTACCTCGACGAAAAG 1191
 Db 266 ATATTTTAGCTATCTGGCGTCTCTTTCAATTCGTATCAGCACTACCTCGACGAAAAG 207
 Qy 1192 AAGTTGAAGACTAGGCACACAAGTCTCTAGGAGATCCTCATCTACTACTACCTGCTGAGGAA 1251
 Db 206 AAGTTGAATACTAGGCACACAAGTCTCTAGGAGATCTTCAAT---CTACCTGTTGAGGAA 150
 Qy 1252 GGTCAACAACATACATAGTGTATGATCAAAAACATATTTCC 1293
 Db 149 GGTCAACAACATACAAAGTGTGATCAAAAGCATATTTCC 108

RESULT 15
 CO105310/c
 LOCUS GR_Eb0035G15.f GR_Eb Gossypium raimondii cDNA clone GR_Eb0035G15
 DEFINITION 5' mRNA sequence.

ACCESSION CO105310 GI:48803996
 VERSION CO105310.1
 KEYWORDS Gossypium raimondii
 SOURCE Gossypium raimondii

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE 1 (bases 1 to 782)
 AUTHORS Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
 Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
 Wing,R.A.

TITLE Global assembly of Cotton ESTs
 JOURNAL Unpublished (2004)
 COMMENT Contact: Rod A. Wing
 Arizona Genomics Institute
 The University of Arizona
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: http://genome.arizona.edu
 Plate: 0035 row: G column: 15.

FEATURES
 source
 1..782
 /organism="Gossypium raimondii"
 /mol_type="mRNA"
 /db_xref="taxon:29730"
 /clone="GR_Eb0035G15"
 /tissue_type="floral"
 /dev_stage="3 to +3 DPA"
 /lab_host="DH10B"
 /clone_lib="GR_Eb"
 /note="Vector: pCMV.SPORT-6.1; Site1: NotI; Site 2:
 EcoRV; Library made by Invitrogen with RNA supplied by
 Wendle lab. Directional cloned into NotI-EV. Colonies
 plated/picked by AGI. More glycerol clones held in -80."

ORIGIN
 Query Match 19.5%; Score 252.2; DB 7; Length 782;
 Best Local Similarity 63.4%; Pred. No. 2.4e-59;
 Matches 386; Conservative 0; Mismatches 223; Indels 0; Gaps 0;

Qy 597 CTCGAGAGTTCACACCTTTTCATTTGTAATCTCTGTTTCTCCTTACTATTTCCTGCG 656
 Db 782 CTCACAAAAGTTCACCCCTTTTCATAATAAACTCTTTGGTCTCTCCATCTCTTCCAC 723
 Qy 657 CCTCTCTGTTGTCACACTGATTCGGAACACACAGAAAAGTGTCTAGAGTAAATAATGT 716
 Db 722 CCTCTCTGTTTCCAAAAGATTCCTCAGGTTCCACACAGTTCACAGAGCCAAATATGT 663
 Qy 717 GATAGGGATAATATACCATTTGCTTCTGCTGGGATTTGATTTGCTATCCCTGGT 776
 Db 662 GATCGGATTTGTATGACGGTGGCTTTCAGACGGGTACGGATTTGATGTTGTCTCTAAC 603

Qy 777 ACAACTGATCCTCAGGAAGGTTTTAAGAAGCAAAACATTTCTCAACGGTCACTGACTTGGT 836
 Db 602 ACAACTTTGTTTCCAAAAGATTTCTCAAGAAAACATCATTTAAGCAGTGTGGACATGAT 543
 Qy 837 CGTTTACCAATCTCTAGTTGCAAGCTGTGTGGTTCTCATAGGACTTTTTCGCAAGCGGGA 896
 Db 542 AATCTACCAAGTCATCGCTGCAACTTTTGATGATCAGTGTGGTCTTTTTCGCCAGTGGGA 483
 Qy 897 GTGAAAACTTTTAAACAAGTGAGATGGAAAACTACAACTGGGAAAAGTGCATACGTTAT 956
 Db 482 ATGGAAGAGTCTGGGCGAAGAAATGGAAGGGTTTAAACTGGGTAAGTCGGCATACTGTTAA 423
 Qy 957 GACTTTGGCCTCGATAGCTATTTCTGGGCAAGTCTACACCATTTGGCGTCTGGGACTGAT 1016
 Db 422 CGTTTGGTTGGATTCTGTAGGGTGGCAAGTTTCTCGATTGGTGGTGGGTTTGAT 363
 Qy 1017 CTTTGGTTCATCTTCTGTGTTTCTCCAAATTCATAACTGTCTGTGGGATTTGCTATPAGTTCC 1076
 Db 362 CTTCGAAGCGTCATCGCTTTTCTCAAATGTGATCAGCACTGTGGACTACCCATTGTTCC 303
 Qy 1077 AGTTGTAGCAAGTGAATTTGTTTCCATGATAAAATGAACGCGTCAAAGATCTTCTCCATCAT 1136
 Db 302 AGTGTTCGATGGTGTGTTTTCACCGATCCCAATGACTGGAATTAAGGTCAATTTCCATGTT 243
 Qy 1137 TTTAGCTATCTGGGGATTTCATTTTCATTTGTCATCAGCACTACTCGACGAAAAAGAGTT 1196
 Db 242 GTTGGCCATATGGGGAATTTGTGCTTAAGTTTACCAGCACTATCTTGATGATCGAAAACTC 183
 Qy 1197 GAAGACTAG 1205
 Db 182 TGATACTGG 174

Search completed: November 1, 2004, 22:30:18
 Job time : 4611.44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:44:43 ; Search time 5291.96 Seconds
(without alignments)
10669.751 Million cell updates

Title: US-09-913-767-5
Perfect score: 1194
Sequence: 1 tcatgagatataataaacat.....tgtcatcaagcatatttcc 1194

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

```

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Query			DB	ID	Description			
	Score	Match	Length						
1	1190.8	99.7	1194	6	AX033548	Sequence			
2	1190.8	99.7	82597	8	ATT9A21	AX021713 Arabidops			
3	1190.8	99.7	194143	8	ATCHRIV48	AL161548 Arabidops			
4	1015.8	85.1	3387	6	AX507517	Sequence			
5	1015.8	85.1	3387	6	AX652015	Sequence			
6	917.2	76.8	1165	8	AY096558	Arabidops			
7	917.2	76.8	1428	8	AY074546	Arabidops			
8	813.4	68.1	1293	6	AX033547	Sequence			
9	813.4	68.1	58427	8	ATF1505	AL110123 Arabidops			
10	570	47.7	733	8	AF083706	Arabidops			
c	11	546.2	45.7	721	11	BX664674 Arabidops			
	12	413	34.6	1106	8	AF370622 Arabidops			
	13	405.8	34.0	1071	6	AX033550 Sequence			
	14	369.6	31.0	1047	6	AX412868 Sequence			
C	15	369.6	31.0	1081	6	AX033549 Sequence			
	16	365.8	30.6	73179	8	AP006383 Lotus cor			
	17	317.2	26.6	1145	6	AX033546 Sequence			
	18	301.6	25.3	110036	8	AC142095 Medicago			
	19	279	23.4	1476	8	AK071254 Oryza sat			

20	279	23.4	1617	8	AK072660	Oryza sat
21	279	23.4	175645	8	AP004096	Oryza sat
22	270	22.6	1251	8	BT003358	Arabidops
23	270	22.6	1372	8	AY062782	Arabidops
24	270	22.6	115421	8	T12C22	Sequence
25	244.2	20.5	1486	8	AK119483	Arabidops
26	244.2	20.5	133299	8	OSJUN0098	Oryza sat
27	244.2	20.5	145739	8	OSJUN00157	Oryza sat
28	215.4	18.0	83698	6	AX540653	Sequence
29	215.4	18.0	83698	8	AB010072	Arabidops
30	213.2	17.9	2175	6	AX654201	Sequence
31	212.8	17.8	1152	6	AX653666	Sequence
32	211.6	17.7	1279	8	BT014339	Lycopersi
33	196	16.4	1117	8	AY096468	Arabidops
34	196	16.4	1333	8	AY074535	Arabidops
35	196	16.4	1403	8	AY087844	Arabidops
36	196	16.4	104386	8	ATY32A17	Arabidops
37	196	16.4	199382	8	ATCHRIV24	Arabidops
38	167.4	14.0	1361	8	AK071680	Oryza sat
39	136.2	11.4	1067	6	AX033553	Sequence
40	135.4	11.3	1049	6	AX033545	Sequence
41	134.4	11.3	1155	6	AX653139	Sequence
42	134.4	11.3	1599	8	AK060241	Oryza sat
43	112.6	9.4	1102	8	BT005504	Arabidops
44	112.6	9.4	1250	8	AK117664	Arabidops
45	111.8	9.4	1225	6	AX033544	Sequence

ALIGNMENTS

RESULT 1

AX033548

LOCUS

Sequence 5 from Patent WO0049152.

ACCESSION

AX033548

VERSION

AX033548.1 GI:10280292

KEYWORDS

. Arabidopsis thaliana (thale cress)

SOURCE

Arabidopsis thaliana

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsia.

REFERENCE

1 Andre,B., Buerkle,L., Frommer,W.B. and Gillissen,B.
Nucleic acids that code for a nucleobase transporter
Patent: WO 0049152-A 5 24-AUG-2000;
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ; GILLISSSEN BERND (DE)
Location/Qualifiers
1..1194
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

FEATURES

source

ORIGIN

Query Match 99.7%; Score 1190.8; DB 6; Length 1194;
Best Local Similarity 99.8%; Pred.No. 0;
Matches 1192; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCATGAGATATAATAAACATGAGTGTTAAATTTTCAGGTGACCAGAGAATTAGGAAGCAAC 60
Db 1 TCATGAGATATAATAAACATGAGTGTTAAATTTTCAGGTGACCAGAGAATTAGGAAGCAAC 60

Qy 61 CTTATGATCATGAGTGGTAACCTGAATCATCATCATCAGCTGTGCCTCAAACCGAGAAC 120
Db 61 CTTATGATCATGAGTGGTAACTGAATCATCATCATCAGCTGTGCCTCAAACCGAGAAC 120

Qy 121 TATAAAGGTGGCTTCGTCTCCATATACGTAATCTTTGTCTCTCTTTGCCAGGCACCTA 180
Db 121 TATAAAGGTGGCTTCGTCTCCATATACGTAATCTTTGTCTCTCTTTGCCAGGCACCTA 180

Qy 181 GCTACAATTCCTGGTAGATTGTACTATGAAAATGGAGGAAATAGCACATATGTGGTAACA 240

Db	181	GCTACAAATCTGGGTAGATTGTACTATGAAAATGGAGAAATAGACACATATGCGTAAACA	240
Qy	241	CTTCTTCAACTCATTTGGCTTCCCTGTACTGGTCTGTGTTCCGCTTCTTTTCTCGAATCAGG	300
Db	241	CTTCTTCAACTCATTTGGCTTCCCTGTACTGGTCTGTGTTCCGCTTCTTTTCTCGAATCAGG	300
Qy	301	CAACCCAAATCAACAGATACAAAATTTAGTCAGTACGTCCTTCCCTTCCACACCCCTTGCATCG	360
Db	301	CAACCCAAATCAACAGATACAAAATTTAGTCAGTACGTCCTTCCCTTCCACACCCCTTGCATCG	360
Qy	361	GTTTACTTGTGCACGTGACGCTAGTGTCCGCTTATGCTTATTTGTCTGCAGTAGGGTTG	420
Db	361	GTTTACTTGTGCACGTGACGCTAGTGTCCGCTTATGCTTATTTGTCTGCAGTAGGGTTG	420
Qy	421	CTCTACTTACAGTCTCTACTTTCTCCCTCATCTTTGGCCTCACAGTTGGCCTTCACTGCC	480
Db	421	CTCTACTTACAGTCTCTACTTTCTCCCTCATCTTTGGCCTCACAGTTGGCCTTCACTGCC	480
Qy	481	TTTTTCTCATATTTCTTAACTCGMAAAGTTCACTCTTTTGATAGTCAGTTCTTTGCTT	540
Db	481	TTTTTCTCATATTTCTTAACTCGMAAAGTTCACTCTTTTGATAGTCAGTTCTTTGCTT	540
Qy	541	CTCCTCACTGATCTCTGCTCTTCTTCTGTGTCACACTGATTCAGAAACTCAACTAAT	600
Db	541	CTCCTCACTGATCTCTGCTCTTCTTGTGTCACACTGATTCAGAAACTCAACTAAT	600
Qy	601	GTATCTAGAGTACAGTATGTGATCGGGTTCATATGTACCATCGGTCTTCGCTGGGATT	660
Db	601	GTATCTAGAGTACAGTATGTGATCGGGTTCATATGTACCATCGGTCTTCGCTGGGATT	660
Qy	661	GGACTGTTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTTTCAGGAACATACATCC	720
Db	661	GGACTGTTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTTTCAGGAACATACATCC	720
Qy	721	TCAGCAGTCAAGGACTTCGGCATTTACAGTCTCTAGTTGGAGTTGTAGTTCTCAT	780
Db	721	TCAGCAGTCAAGGACTTCGGCATTTACAGTCTCTAGTTGGAGTTGTAGTTCTCAT	780
Qy	781	GGACTTTTTCAGAGTGGAGTGGAACTTTGCCAAGTGAGATGAGAAACTACAACTC	840
Db	781	GGACTTTTTCAGAGTGGAGTGGAACTTTGCCAAGTGAGATGAGAAACTACAACTC	840
Qy	841	GGGAAAGTGCATATGTTTGACTTTAGCTTCGGCAGCTATTTCTGCGAAGTCTACACT	900
Db	841	GGGAAAGTGCATATGTTTGACTTTAGCTTCGGCAGCTATTTCTGCGAAGTCTACACT	900
Qy	901	CCTGCTCTTGTGGGATTGATCTTCAGTGCATCTCTGTTTCTCAATTCCATAACAGCT	960
Db	901	CCTGCTCTTGTGGGATTGATCTTCAGTGCATCTCTGTTTCTCAATTCCATAACAGCT	960
Qy	961	GTGGGATTCSCCTATAGTTCCAGTTTCGGCAGTGTAGTTTTCATGATAGAAATGGACGCA	1020
Db	961	GTGGGATTCSCCTATAGTTTCAGTTTCGGCAGTGTAGTTTTCATGATAGAAATGGACGCA	1020
Qy	1021	TCCAAAATCTTCTCCATTAATTTTAGCTATCTGCGGCTTCCTTTCATTCGTCATCAGCAC	1080
Db	1021	TCCAAAATCTTCTCCATTAATTTTAGCTATCTGCGGCTTCCTTTCATTCGTCATCAGCAC	1080
Qy	1081	TACCTCGACGAAAGAGTGAATTAAGTACTAGCCACACAGTGTGTAGGAGATCTTTCATCTA	1140
Db	1081	TACCTCGACGAAAGAGTGAATTAAGTACTAGCCACACAGTGTGTAGGAGATCTTTCATCTA	1140
Qy	1141	CCTGTTGAGGAGGTTCACAAACATACAAAGTGTGTGATCAAAAGCATATTTC	1194
Db	1141	CCTGTTGAGGAGGTTCACAAACATACAAAGTGTGTGATCAAAAGCATATTTC	1194

RESULT 2
ATT9A21
LOCUS
DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone T9A21 (ESSA project).

ACCESSION	AL021713
VERSION	AL021713.1
KEYWORDS	GI:2832689
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 Bevan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H.W., Mayer, K.P.X., Lemcke, K. and Schueller, C.
AUTHORS	Unpublished
JOURNAL	2 (bases 1 to 82697)
REFERENCE	EU Arabidopsis sequencing, project.
AUTHORS	Direct Submission
TITLE	Submitted (22-SEP-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schueller@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ .
FEATURES	Location/Qualifiers
source	1..82697
	/organism="Arabidopsis thaliana"
	/mol_type="genomic DNA"
	/variety="Columbia"
	/db_xref="taxon:3702"
	/chromosome="4"
gene	402..1585
	/gene="T9A21.10"
CDS	join(402..832,946..1090,1178..1585)
	/gene="T9A21.10"
	/notes="similarity to SPF1 protein, sweet protein, PIR2:551529 and WRKY protein family, Petroselinum crispum, MNOS:S72443, MNOS:S72444, MNOS:S72445"
	/codon_start=1
	/product="DNA binding-like protein"
	/protein_id="CAA16788.1"
	/db_xref="GI:4038481"
	/db_xref="GOA:O49720"
	/db_xref="GOA:Q8VWJ2"
	/db_xref="Swiss-Prot:Q8VWJ2"
	/translation="MSNETRLNYQYPPSSFSILHEMNLPTSPSSVGNLPQNGFNP STYFTDCIQSSPAAYELDLQTFGLSPSSSEVFNSSIDQEPNRDVTNDVINGACNE TETRVGPSNSSSEADHPGDSGKRRKRELVEEQIISKVGKTKTTEVKKQREPRV SFTWKSEVDHLRDGYRWKYGOKAVKNSPYPIIANGNENRYYRCTKCNVKRVE RSQDPTVITTYEGQHNHPIPTNLGSSSAAAFSADLMTPRSFAHDMFTAAATNG GSVAALDYGQSGYGSVNSNPSSHQVHQGEYELLREIFFSIFPKQEP"
exon	402..832
	/gene="T9A21.10"
	/number=1
intron	833..945
	/gene="T9A21.10"
	/number=1
exon	946..1090
	/gene="T9A21.10"
	/number=2
intron	1091..1177
	/gene="T9A21.10"
	/number=2
exon	1178..1585
	/gene="T9A21.10"
	/number=3
gene	4778..6211
	/gene="T9A21.20"
CDS	join(<4778..4921,5023..5187,5263..5552,5638..5866,5951..6211)
	/gene="T9A21.20"
	/notes="strong similarity to exopolysgalacturonase, Arabidopsis thaliana, PIR2:S34266"

Db	10985	CTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTGTTCCGCTCTTTCTCGAATCAGG	11044
Qy	301	CAACCCAAATCAACAGATACAAAATTTTCAGTCAGTCCCTTCCCTCACACACCTTGCATCG	360
Db	11045	CAACCCAAATCAACAGATACAAAATTTTCAGTCAGTCCCTTCCCTCACACACCTTGCATCG	11104
Qy	361	GTTTACTTGTGCACTGGACTGCTAGTGTCCGCTTATGCTTATTTGTCGAGTAGGGTTG	420
Db	11105	GTTTACTTGTGCACTGGACTGCTAGTGTCCGCTTATGCTTATTTGTCGAGTAGGGTTG	11164
Qy	421	CTCTACTTACCAGTCTCTACTTCTTCCCTCATCTTGGCTCACAGTTGGCTTCACTGCC	480
Db	11165	CTCTACTTACCAGTCTCTACTTCTTCCCTCATCTTGGCTCACAGTTGGCTTCACTGCC	11224
Qy	481	TTTTTCTCATATTTCTTTAACTCGCAAAAGTTCACTCCTTTTGATAGTCAGTCTTTGCTT	540
Db	11225	TTTTTCTCATATTTCTTTAACTCGCAAAAGTTCACTCCTTTTGATAGTCAGTCTTTGCTT	11284
Qy	541	CTCCTCACTGTATCTCTGCTCTTCTTGTGTGCTCAACACTGATTCAGAAAACTCAACTAAT	600
Db	11285	CTCCTCACTGTATCTCTGCTCTTCTTGTGTGCTCAACACTGATTCAGAAAACTCAACTAAT	11344
Qy	601	GTATCTAGTACAGTATGTGATCGGTTTCATATGTACCATCGGTGCTTCGCTGGGATT	660
Db	11345	GTATCTAGTACAGTATGTGATCGGTTTCATATGTACCATCGGTGCTTCGCTGGGATT	11404
Qy	661	GGACTGTTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTTTCAGAAAGCATACATCC	720
Db	11405	GGACTGTTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTTTCAGAAAGCATACATCC	11464
Qy	721	TCAGCAGTCAAGGCTTGGCCATTTACAGTCTCTAGTTGGAGTTGTAGTTCTCATTA	780
Db	11465	TCAGCAGTCAAGGCTTGGCCATTTACAGTCTCTAGTTGGAGTTGTAGTTCTCATTA	11524
Qy	781	GGACTTTTTCGAAGTGGAGAGTGGGAAACTTTGCCAGTGAGTACAGAACTACAAATC	840
Db	11525	GGACTTTTTCGAAGTGGAGAGTGGGAAACTTTGCCAGTGAGTACAGAACTACAAATC	11584
Qy	841	GGGAAAGTGTATATGTTTTCACCTTTAGCTTCGGCAGCTATTTCTCGCAAGTCTACACT	900
Db	11585	GGGAAAGTGTATATGTTTTCACCTTTAGCTTCGGCAGCTATTTCTCGCAAGTCTACACT	11644
Qy	901	CCTGGTCTTGTGGATTCATCTTCAGAGTCATCCTCTGTGTTCTCCAATTCATTAACAGCT	960
Db	11645	CCTGGTCTTGTGGATTCATCTTCAGAGTCATCCTCTGTGTTCTCCAATTCATTAACAGCT	11704
Qy	961	GTGGGATTCCTATAGTTCCAGTTTCAGGTCAGTATGTTTCCATGATAGAAATGAGCGCA	1020
Db	11705	GTGGGATTCCTATAGTTCCAGTTTCAGGTCAGTATGTTTCCATGATAGAAATGAGCGCA	11764
Qy	1021	TCCAAATCTTCTCCATTTATTTAGCTATCTCGCGCTTCCTTTTCATTGCTATCAGCAC	1080
Db	11765	TCCAAATCTTCTCCATTTATTTAGCTATCTCGCGCTTCCTTTTCATTGCTATCAGCAC	11824
Qy	1081	TACCTCGAGAAAGAGTGAATGAACTAGCCACACAGTCTGTAGGAGATCTTCACTTA	1140
Db	11825	TACCTCGAGAAAGAGTGAATGAACTAGCCACACAGTCTGTAGGAGATCTTCACTTA	11884
Qy	1141	CCTGTTGAGGAAGGTTCACAAAAACATCAAAAGTGTGTCATCAAAAGCATATTTC	1194
Db	11895	CCTGTTGAGGAAGGTTCACAAAAACATCAAAAGTGTGTCATCAAAAGCATATTTC	11938

RESULT 3
ATCHRIV48
LOCUS ATCHRIV48 194143 bp DNA linear PLN 16-MAR-2000
DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 48.
ACCESSION AL161548
VERSION AL161548.2 GI:7268504
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana

Eukaryota: Viridiplantae: Streptophyta; Embryophyta: Tracheophyta; Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots, rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 111084)
Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
Unpublished
2 (bases 107966 to 194143)
Halbert,H., Braun,M., Hoizer,E., Brandt,A., Duesterhoeft,A., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
Unpublished
3 (bases 1 to 194143)
EU Arabidopsis sequencing, project.
Direct Submission
Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de
Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>
this fragment has an overlap with ATCHRIV47 at the 5' end and an overlap with ATCHRIV49 at the 3' end.
FEATURES
Location/Qualifiers
1..194143
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
3377..9293
/gene="AT4g18120"
join(6377..6796,7122..7238,7347..7501,7578..8456,8543..8627,8716..8784,8871..9293)
/gene="AT4g18120"
/note="similarity to Me12-like protein, Arabidopsis thaliana, gb:D86122 contains EST gb:W43784, AI998063.1"
/codon_start=1
/product="putative protein"
/protein_id="CA78814.1"
/db_xref="GI:7268605"
/db_xref="GOA:Q9SVV9"
/db_xref="TrEMBL:Q9SVV9"
/translation="MASGLDHFSGIGNMLDDGDSHPIGNMLPDEEELFSGMLDDLNLSSLPATLDLEDYDLFGSGGLEETDPYDSLNGKGFSRMGFADSNVDNTPQNIQFNGVSIAGEHPYGBHPSTLTFVRNINSNVEDSELOALPEDNPSEKDVNQGLTVFNLAPSVNRDLNITFGVYGEIKQIRTPNKRHHKFEFTDVRSDAALKALNRTEIAGKRIKLHSPRGGARRNNMLQMLEDDDSYSLNHVESPLASSPIGNWRNSPIDPLQSPSKSPIGNLSPTNIRYFESMKTASVNDQGRFRFSLDHLDFSSSYNNASHKASTPQQPSFGVSFGLNSHPSHVETLSGSEFLWGSFSSAWPNPFPSSNRNHRFPYSAQNSGLQHHIGAPSGFFPRSPETSMGSAFRGAGNMAQRNLRETSSPNFKMLSAPRSOLFTGSGYLPAAATMSVIDDPLEGDSNQPDNSGNQADIKTFOLDLSKIMRGEPRPTLKNTPNKYTRNNMLAAIDENKSCYDLYLPIDENKKNVGYAFKNMVSFKETIAYEAFNKKDKFKSEKSLAYATQKKAALIAHFQNSSLMNSLMDRRCCQPLVFDSESKYPNNLLQTHDLKMIILYFCVYLMNWLTHTCUFETRSGSEFAARASNSTGFLIFLGARTHIKADSQNVNHHVIAISGLKRKYQFLG"
6377..6796
/gene="AT4g18120"
/number=1
6797..7121
/gene="AT4g18120"
/number=1
7122..7238
/gene="AT4g18120"
/number=2
7239..7346
/gene="AT4g18120"
/number=2
7347..7501
/gene="AT4g18120"

|||||
Db 39432 CAACCCAAATCAACAGATACAAATTCAGTCAGTCCCTCTTCCACACCCCTTGCATCG 39491
Qy 361 GTTTACTGTGCACTGGACCTGCTAGTGCCTTATGCTTATTTGTCTGAGTAGGGTTG 420
Db 39492 GTTTACTGTGCACTGGACCTGCTAGTGCCTTATGCTTATTTGTCTGCACTAGGGTTG 39551
Qy 421 CTCTACTTACCAGTCTCTACTTTCTCCCTCATCTTGGCCCTCACAGTTGGCTTCACATGCC 480
Db 39552 CTCTACTTACCAGTCTCTACTTTCTCCCTCATCTTGGCCCTCACAGTTGGCCCTTCACATGCC 39611
Qy 481 TTTTCTCATATTTCCCTTAACCTCGCAAAAGTTCACTCTCTTGATAGTCAGTTCTTTGCTT 540
Db 39612 TTTTCTCATATTTCCCTTAACCTCGCAAAAGTTCACTCTCTTGATAGTCAGTTCTTTGCTT 39671
Qy 541 CTCCTCACTGATCCTCTGCTCTTCTTGTGCTCAACACTGATTCAGAAAACCTCAACTAAT 600
Db 39672 CTCCTCACTGATCCTCTGCTCTTCTTGTGCTCAACACTGATTCAGAAAACCTCAACTAAT 39731
Qy 601 GTATCTAGAGTACAGTATGATCGGGTTCATATGTACCATCGTGTCCGCTGGGATT 660
Db 39732 GTATCTAGAGTACAGTATGATCGGGTTCATATGTACCATCGTGTCTCCGCTGGGATT 39791
Qy 661 GGACTGTCTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTTTCAGCAAGCATACATCC 720
Db 39792 GGACTGTCTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTTTCAGCAAGCATACATCC 39851
Qy 721 TCAGCACTGACCGACTGGCCATTTCACAGTCTCTAGTTCGAGTGTGTAGTTCTCAT 780
Db 39852 TCAGCACTGACCGACTGGCCATTTCACAGTCTCTAGTTCGAGTGTGTAGTTCTCAT 39911
Qy 781 GGACTTTTTCAGTGGAGTGGGAACTTTTGCAGAGTGAGATGAGAACTACAAACTC 840
Db 39912 GGACTTTTTCAGTGGAGTGGGAACTTTTGCAGAGTGAGATGAGAACTACAAACTC 39971
Qy 841 GGGAAAGTGTCAATGTTTTCAGTTCCTCGGAGCTATTTCTCGCAAGTCTACACT 900
Db 39972 GGGAAAGTGTCAATGTTTTCAGTTCCTCGGAGCTATTTCTCGCAAGTCTACACT 40031
Qy 901 CCTGGCTGTGGGATTGATCTTCAGATCACTCTCTGTGTCTCAATTCATTAACAGCT 960
Db 40032 CTTCGGTCTGTGGGATTGATCTTCAGATCACTCTCTGTGTCTCAATTCATTAACAGCT 40091
Qy 961 GTGGGATGTCCTATAGTTCAGTTCGCGAGTGTATTTTCCATGATAGATGAGCGCA 1020
Db 40092 GTGGGATGTCCTATAGTTCGAGTTCGCGAGTGTATTTTCCATGATAGATGAGCGCA 40151
Qy 1021 TCCAAATCTTCTCCATTTATTTAGCTATCTCGGCTTCCCTTTCATTCGCTCATACAGCAC 1080
Db 40152 TCCAAATCTTCTCCATTTATTTAGCTATCTCGGCTTCCCTTTCATTCGCTCATACAGCAC 40211
Qy 1081 TACCTCGAGAAAAGAGTTGAATACCTAGCCACACAAAGTGTGTAGGAGATCTTTCATCTA 1140
Db 40212 TACCTCGAGAAAAGAGTTGAATACCTAGCCACACAAAGTGTGTAGGAGATCTTTCATCTA 40271
Qy 1141 CCTGTTGAGGAGGTCACACAAATACAAAGTGTGTATCAAGCATATTTCC 1194
Db 40272 CCTGTTGAGGAGGTCACACAAATACAAAGTGTGTATCAAGCATATTTCC 40325

RESULT 4
AX507517 AX507517 3387 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 2212 from Patent WO0216655.
ACCESSION AX507517
VERSION AX507517.1 GI:23388754

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1

AUTHORS Harper, J. F., Kzepe, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing same, and methods of use
JOURNAL Patent: WO 0216655-A 2212 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG (CH)
FEATURES Location/Qualifiers
source 1. 3387
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 85.1%; Score 1015.8; DB 6; Length 3387;
Best Local Similarity 95.1%; Pred. No. 3.1e-273; Indels 54; Gaps 1;
Matches 1081; Conservative 0; Mismatches 2;
Qy 38 GTGACCAGAACTTAGAAGCAAACTTATAGATCATGAGGTGGTAACTGAATCATCATCAT 97
Db 1205 GTGACCAGAACTTAGAAGCAAACTTATAGATCATGAGGTGGTAACTGAATCATCATCAT 1264
Qy 98 CAGCTGTGCTCAAAACCGAGAACTATAAAAGGTGGCTTCGTCTCCATATACGTAATCT 157
Db 1265 CAGCTGTGCTCAAAACCGAGAACTATAAAAGGTGGCTTCGTCTCCATATACGTAATCT 1324
Qy 158 TTGTCTCTTTTCCAGCCACTAGCTACAAATCTGGGTAGATTTGTACTATGAAATCGAG 217
Db 1325 TTGTCTCTTTTCCAGCCACTAGCTACAAATCTGGGTAGATTTGTACTATGAAATCGAG 1384
Qy 218 GAAATAGCACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTGT 277
Db 1385 GAAATAGCACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTTCTGT 1444
Qy 278 TCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCCC 337
Db 1445 TCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCCC 1504
Qy 338 CTTCCTTCCACACCTTGCATCGGTTTACTTGTGCACCTGCTAGTGTCCGCTTATG 397
Db 1505 CTTCCTTCCACACCTTGCATCGGTTTACTTGTGCACCTGCTAGTGTCCGCTTATG 1564
Qy 398 CTATTTTGTCTGCACTAGGTTGCTCTACTTACCACTCTCTCTTCTCCCTCATCTGG 457
Db 1565 CTATTTTGTCTGCACTAGGTTGCTCTACTTACCACTCTCTCTTCTCCCTCATCTGG 1578
Qy 458 CTTCACAGTTGGCTTCACTGCTTTTCTCATATTTTCTTAACTCGCAAAAGTCACTC 517
Db 1579 -----TTGGGCTTCACTGCTTTTCTCATATTTTCTTAACTCGCAAAAGTCACTC 1630
Qy 518 CTTTGATAGTCAGTTCTTCTCTCTCACTGCTATCTCTGCTCTCTTCTTGTGGTCAACA 577
Db 1631 CTTTGATAGTCAGTTCTTCTCTCTCTCACTGCTATCTCTGCTCTCTTCTTGTGGTCAACA 1690
Qy 578 CTGATTCAGAAAACTCAACTAATGATCTAGAGTACAGTATGTGATCGGGTTCATATGTA 637
Db 1691 CTGATTCAGAAAACTCAACTAATGATCTAGAGTACAGTATGTGATCGGGTTCATATGTA 1750
Qy 638 CCATCGGTGCTTCGCTGGGATTTGGAATGTTACTCTCTGATACAAATGCTCTTCAGGA 697
Db 1751 CCATCGGTGCTTCGCTGGGATTTGGAATGTTACTCTCTGATACAAATGCTCTTCAGGA 1810
Qy 698 AAGTTTTTCAGAAAGCATACATCTCAGCAGTCAGGACTTGGCCATTTACAGTCTCTAG 757
Db 1811 AAGTTTTTCAGAAAGCATACATCTCAGCAGTCAGGACTTGGCCATTTACAGTCTCTAG 1870
Qy 758 TTGCGAGTTGTAGTTCTCATAGGACTTTTTCGAGTGGAGAGTGGGAACTTTGCGCAA 817
Db 1871 TTGCGAGTTGTAGTTCTCATAGGACTTTTTCGAGTGGAGAGTGGGAACTTTGCGCAA 1930
Qy 818 GTGAGATGAGAACTACAACTCGGGAAGTGTCTATGTTTGGACTTTAGCCTTCGGCAG 877
Db 1931 GTGAGATGAGAACTACAACTCGGGAAGTGTCTATGTTTGGACTTTAGCCTTCGGCAG 1990

```
QY 878 CTAATTTCTGGGCAAGTCTACACCTCTGGTCTTGTGGGATTGATCTTCGAGTATCCTCTCG 937
Db 1991 CTAATTTCTGGGCAAGTCTACACCTCTGGTCTTGTGGGATTGATCTTCGAGTATCCTCTCG 2050
QY 938 TGTTCCTCCAATCCATAACAGCTGTGGGATTGGCTATAGTTCAGATTGGCGGAGTGATAG 997
Db 2051 TGTTCCTCCAATCCATAACAGCTGTGGGATTGGCTATAGTTCAGATTGGCGGAGTGATAG 2110
QY 998 TTTTCCATGATAGATGGAGCATTCCAAAATCTTCTCCATTATTTTATAGTATCTCGGGCT 1057
Db 2111 TTTTCCATGATAGATGGAGCATTCCAAAATCTTCTCCATTATTTTATAGTATCTCGGGCT 2170
QY 1058 TCTTTTCATCTGCTATCAGCACTACCTCGAGAAAGATTGAATATAGCCACAA 1117
Db 2171 TCTTTTCATCTGCTATCAGCACTACCTCGAGAAAGATTGAATATAGCCACAA 2230
QY 1118 GTGCTGTAGGAGATCTTCACTACCTGTTGAGGAAGTCCACAAACATACAAAGTG 1174
Db 2231 GTGCTGTAGGAGATCTTCACTACCTGTTGAGGAAGTCCACAAACATACAAAGTG 2287

RESULT 5
AX652015 3387 bp DNA linear PAT 22-MAR-2003
LOCUS Sequence 908 from Patent WO0300899.
DEFINITION AX652015
ACCESSION AX652015
VERSION AX652015.1 GI:29154833
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Chang, H. S., Chen, W., Cooper, B., Glazebrook, J., Goff, S. A., Hou, Y. M.,
Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, F. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 0300898-A 908 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
Source
1. 3387
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 85.1%; Score 1015.8; DB 6; Length 3387;
Best Local Similarity 95.1%; Pred. No. 31e-273;
Matches 1081; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

QY 38 GTGACCAAGACTTAGAAGCAACCTTATAGATCATGAGTGGTAACTGAATCATCATCAT 97
Db 1205 GTGACCAAGACTTAGAAGCAACCTTATAGATCATGAGTGGTAACTGAATCATCATCAT 1264
QY 98 CAGCTGTGCTCAAAACCGAGAACTATAAAAGTGGCTTCGTCTCCATATACGTAATCT 157
Db 1265 CAGCTGTGCTCAAAACCGAGAACTATAAAAGTGGCTTCGTCTCCATATACGTAATCT 1324
QY 158 TTTGCTCTTTTGGCCAGCCACTAGCTACAAATCTGGGTAGATTGACTATGAATGGAG 217
Db 1325 TTTGCTCTTTTGGCCAGCCACTAGCTACAAATCTGGGTAGATTGACTATGAATGGAG 1384
QY 218 GAATAGACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGACTGTTCTGT 277
Db 1385 GAATAGACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGACTGTTCTGT 1444
QY 278 TCGGCTTCTTTCTGAATCAGGCAACCCAAATCAACAGATCAAAATTCAGTCAGTCCC 337
Db 1445 TCGGCTTCTTTCTGAATCAGGCAACCCAAATCAACAGATCAAAATTCAGTCAGTCCC 1504
QY 338 CTTCTCTACACCCCTTGATCGGTTTACTTGTGCACTGGACTGCTAGTTCGGTTATG 397
Db 1505 CTTCTCTACACCCCTTGATCGGTTTACTTGTGCACTGGACTGCTAGTTCGGTTATG 1564
```

```
QY 398 CTTATTTGTCTGCAGTAGGGTTGCTCTACTTACCAGTCTCTACTTTCTCCCTCATCTTGG 457
Db 1565 CTTATTTGTCTGCA----- 1578
QY 458 CCTCAGAGTTGGCCCTTCACTGCTTTTCTCATATTTCTTAATCTCGGAAAGTTCACTTC 517
Db 1579 -----TTGGCCCTTCACTGCTTTTCTCATATTTCTTAATCTCGGAAAGTTCACTTC 1630
QY 518 CTTTGTATGTCAGTTCTTTTGTCTTCCCTCAGTATCCTCTGCTCTTCTTGTGGTCAACA 577
Db 1631 CTTTGTATGTCAGTTCTTTTGTCTTCCCTCAGTATCCTCTGCTCTTCTTGTGGTCAACA 1690
QY 578 CTGATTTCAGAAAACCTCAACTAATGTATCTAGAGTACAGTATGTGATCGGTTTCATATGTA 637
Db 1691 CTGATTTCAGAAAACCTCAACTAATGTATCTAGAGTACAGTATGTGATCGGTTTCATATGTA 1750
QY 638 CCATCGGTGCTTCGCTGGGATTGGACTGTGTACTATCTCTGATACAAATGCTCTTCAGGA 697
Db 1751 CCATCGGTGCTTCGCTGGGATTGGACTGTGTACTATCTCTGATACAAATGCTCTTCAGGA 1810
QY 698 AAGTTTTTCAGGAAGCATACATCCTCAGCAGTCCAGGACTTTGGCCATTTTACCAGTCTCTAG 757
Db 1811 AAGTTTTTCAGGAAGCATACATCCTCAGCAGTCCAGGACTTTGGCCATTTTACCAGTCTCTAG 1870
QY 758 TTGCGAGTTGTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGAGTGGGAAACTTTTGCCAA 817
Db 1871 TTGCGAGTTGTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGAGTGGGAAACTTTTGCCAA 1930
QY 818 GTGAGATGAGAAAACCTACAAAACCTCGGGAAGTGTCTATGTGTTTGTAGCTTTAGGCTCGGAG 877
Db 1931 GTGAGATGAGAAAACCTACAAAACCTCGGGAAGTGTCTATGTGTTTGTAGCTTTAGGCTCGGAG 1990
QY 878 CTAATTTCTGCAAGTCTACACCTCTGCTGCTTGTGGGATTGATCTTCAGTCACTCTCTG 937
Db 1991 CTAATTTCTGCAAGTCTACACCTCTGCTGCTTGTGGGATTGATCTTCAGTCACTCTCTG 2050
QY 938 TGTTCCTCAATTTCCATAACAGCTGTGGGATTGCCCTATAGTTCAGTTGCGGAGTGATAG 997
Db 2051 TGTTCCTCAATTTCCATAACAGCTGTGGGATTGCCCTATAGTTCAGTTGCGGAGTGATAG 2110
QY 998 TTTTCCATGATAGATGGAGCAGTATCCAAAATCTTCTCCATTTATTTAGTATCTCGGGCT 1057
Db 2111 TTTTCCATGATAGATGGAGCAGTATCCAAAATCTTCTCCATTTATTTAGTATCTCGGGCT 2170
QY 1058 TCTTTTCATCTGCTATCAGCACTACCTCGAGAAAGATTGAATGTAATAGTACTAGCCACAA 1117
Db 2171 TCTTTTCATCTGCTATCAGCACTACCTCGAGAAAGATTGAATGTAATAGTACTAGCCACAA 2230
QY 1118 GTGCTGTAGGAGATCTTCACTTACCTGTTGAGGAAGTCCACAAACATACAAAGTG 1174
Db 2231 GTGCTGTAGGAGATCTTCACTTACCTGTTGAGGAAGTCCACAAACATACAAAGTG 2287
```

```
RESULT 6
AY096558 LOCUS
DEFINITION Arabidopsis thaliana putative protein (At4g18200) mRNA, complete cds.
ACCESSION AY096558
VERSION AY096558.1 GI:20465496
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1165)
REFERENCE
AUTHORS Yamada, K., Banh, J., Chan, M. M., Chang, C. H., Chang, E., Dale, J. M.,
Deng, J. M., Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L.,
Tang, C., Toriumi, M., Wu, H. C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
```


SOURCE ORGANISM	Arabidopsis thaliana (thale cress)	
REFERENCE	Arabidopsis thaliana	
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
JOURNAL	1	
REFERENCE	Bevan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H.W., Mayer, K.F.X., Lemcke, K. and Schueller, C.	
AUTHORS	Unpublished	
TITLE	2 (bases 1 to 58427)	
JOURNAL	EU Arabidopsis sequencing, project.	
REFERENCE	Direct Submission	
AUTHORS	Submitted (30-MAR-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schueller@mips.biochem.mpg.de, maye@mips.biochem.mpg.de	
JOURNAL	Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk	
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ .	
FEATURES	Location/Qualifiers	
source	1..58427	
	/organism="Arabidopsis thaliana"	
	/mol_type="genomic DNA"	
	/variety="Columbia"	
	/db_xref="taxon:3702"	
	/chromosome="4"	
misc_feature	1..2548	
	/note="overlap to BAC T6K21; please refer to EMBL:AL021889 for analysis and annotation"	
gene	3596..5023	
gene	/gene="F15J5.10"	
	complement(join(3596..3646,3736..3801,3895..4020,4116..4281,4725..5023))	
CDS	/gene="F15J5.10"	
	complement(join(3596..3646,3736..3801,3895..4020,4116..4281,4725..5023))	
	/gene="F15J5.10"	
	/note="Contains Eukaryotic initiation factor 4E signature [DPYCFKHIEPKWEDPICANGKW]	
	contains EST gb:721435, AAS86111"	
	/codon_start=1	
	/product="translation initiation factor eIF4E"	
	/protein_id="CAB53645.1"	
	/db_xref="GI:5816990"	
	/db_xref="GOA:023252"	
	/db_xref="Swiss-Prot:O23252"	
	/translation="WAVEDTPKSVVTEAKPNISNPIDRYHEGDDAEERGEIAGGEG DGNVDESKSGVPESHLEHWTWFDNPVAVKSKOTSWGSLRPVFTFTSEEFWSLY NNMKHPSLAHGADFCFKHIIPEKWEDPICANGKGTWTFPKSKDSKSWLYTLALI GEQFDHGEICGAVNIRGQKERISITWTKNASNEAAQVSIGKQWKEFLDYNNSIGFII HEDAKLDNKNVTA"	
exon	complement(3596..3646)	
	/gene="F15J5.10"	
intron	/number=1	
	complement(3647..3735)	
exon	/gene="F15J5.10"	
	/number=1	
	complement(3736..3801)	
intron	/gene="F15J5.10"	
	/number=2	
	complement(3802..3894)	
exon	/gene="F15J5.10"	
	/number=2	
	complement(3895..4020)	
intron	/gene="F15J5.10"	
	/number=3	
	complement(4021..4115)	
exon	/gene="F15J5.10"	
	/number=3	
	complement(4116..4281)	
	/gene="F15J5.10"	
	/number=4	

intron		
exon		
repeat_region		
gene		
CDS		
	join(9077..9301,9805..9859,9938..10113,10198..10419,10487..10725,10806..11331,11419..11692,11793..12736,12827..13093,13169..13432,13508..13984)	
	/gene="F15J5.20"	
	/note="similarity to multidrug resistance protein 1, Aspergillus fumigatus, U62933	
	Contains ABC transporters family signature, ABC Transporter [MSGGQRLAIARAI] [LUSGGQRIARAI], ATP/GTP-binding site motif A (P-loop), Atp_Gtp_A [GQSGSGKS] [GESGSGKS]"	
	/codon_start=1	
	/product="multidrug resistance protein/P-glycoprotein-like"	
	/protein_id="CAB53646.1"	
	/db_xref="GI:5816991"	
	/db_xref="GOA:Q9SVW6"	
	/db_xref="TrEMBL:Q9SVW6"	
	/translation="NEEKSSKNNGGNQKVSFFKLPSEADKTDVLMVTGTTAAAGNG LTQPMILIFQLINAFGTTDPDHVRVWVKVFIYLAIVTSVVAFLQVSCWMTG ERQATIRGLYKLTIRQDIDGYFDTETNTVIGRMSGDTILIQDMGERVKFTQLL CTFLGGFAIFKGPLLAGVLCSCITPLIVIAAGASLIMSXMGARGQVAYABAGNVLE QTVGAIRTVVAFTEKQATEKYEKLEIAYKTVVOOGLISGFLGTMALVIFCSYGLA VYGAKLIMEKGYNGGVINVIFAVLTCGMSIGOTSPSLNAPFAAGRAAFKWFETIKR SPKIDAVMSSGVLEDIRGDIELKDVPYRPAPDVOI/PAGSLFVPGTKVALVQQS GSGKTSVISLERFDPPSGQVLIDNLDKQLKIRSKIGLVQSEPPVLFATIKEN LAYKEDATDQEIIRTAELANAAPFDKLPGLDPTWVGEHGTMSGGQKQRLAIARAI LKNPKILLDEATSALESERI/QDVALNLSNRITVVVAHRLTIRTADVIAVHQ GKIVBKGTDMIQDPEGAYSQLVQESKEATESESRPETSRLDVERSGSLRSAM RRSVRSNSSHSHSFLASNNMFFQGVNVDMEDEENNVRHKKVSLKRLAHKNKE IPLVLGSIAMVHGTVPIFGLLSSINMFPYEPAKILKDSHPWALIYALGTLNF VMIPVNTFFGIAGKLIKIRSMCFDKVHQEISWFDOTANSRYNFIYIINRRLY LILIFICVLLPPLVRLRECTDASTVRSVLGDLALIVQNIATVTTGLIITAFITANI LALIVLALSPFIVIQGVAQTFLTGFSAKAMYTEASQVANDAVSSIRTVASFCAEE KVMLDYQOKDGPKNQVRLGLSGAGGFSEFFLYCINCVCFVSGAGLIOIGKATFG EVKVFVFTALTIMAGVSTSAMADSNKAKDASIPDLDSTPKIDSSDEGTTLQNVNGDIEPHUSFRYPMRPDVOI/PDRLCLTIPSGKTVALVSGSGSKTSVSMIERFYN PDSKILLDQVEIOTFKLSLURQOMGLVQSEPIFLFNETIRSNIRATGKTGGATEEIIA AAKAANAHNFISSLPQGYDTSVGERGVQLSGGQKQRIAIARAILKPKILLDEATSA LDAESRVVQDALDRVXG"	
	9077..9301	
	/gene="F15J5.20"	
	/number=1	
	9302..9804	
	/gene="F15J5.20"	
	/number=1	
	9805..9859	
	/gene="F15J5.20"	
	/number=2	
	9860..9937	
	/gene="F15J5.20"	
	/number=2	
	9938..10113	
	/gene="F15J5.20"	
	/number=3	
	10114..10197	
	/gene="F15J5.20"	
	/number=3	
	10198..10419	
	/gene="F15J5.20"	
	/number=4	
	10420..10486	
	/gene="F15J5.20"	
	/number=4	

[illegible]

REFERENCE 1 (bases 1 to 733)
AUTHORS Stracke,R. and Palme,K.
TITLE Signal Peptide Selection derived cDNAs from Arabidopsis thaliana
leaves and guard cells
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 733)
AUTHORS Stracke,R. and Palme,K.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-1998) Max-Deibruueck-Laboratorium in der
Max-Planck-Gesellschaft, Carl-von-Linne-Weg 10, Koeln D-50829,
Germany

FEATURES
source Location/Qualifiers
1..733
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="aps258"
/note="signal peptide selection derived"
38..>733
/codon_start=1
/product="unknown"
/protein_id="AA060265.1"
/db_xref="GI:24417310"
/translation="MGISQVHY/CNGDNLEANLLDHEETSESPVQTXCKRWLRVSI
YAFVFCPLAVLGRLYIENGKSTVYVLLQLIGPVLILFRFFKIRPKSTDT
NFSQPSFTTASVYXCTGLLVSAYILSAVGLLYPVSTFSLIASQLAFAPESYF
LNSQKFTPLIVNSLFLTVSSALLVNTDSENTTNSRVQYVIGFICTIGASAGIGLV
LSLIQLFRKVFVK"

ORIGIN
Query Match 47.7%; Score 570; DB 8; Length 733;
Best Local Similarity 91.3%; Pred. No. 2.5e-148;
Matches 616; Conservative 0; Mismatches 50; Indels 9; Gaps 1;
QY 37 GTGTACCAAGTGAACCAACCTTATAGATCATGAGGTGGTAACTCAATCATCA 96
DB 68 GTGTACCAAGTGAACCAACCTTATAGATCATGAGGTGGTAACTCAATCATCA 118
QY 97 TCAGTGTGCTCAACACGAGAACTATAAAGGTGCTTCGTGCTCCATATACGTAATC 156
DB 119 TTTTCAGTCACTCAACGAGAACTATAAAGGTGCTTCGTGCTCCATATACGTAATC 178
QY 157 TTTTCAGTCACTCAACGAGAACTATAAAGGTGCTTCGTGCTCCATATACGTAATC 216
DB 179 TTTTCAGTCACTCAACGAGAACTATAAAGGTGCTTCGTGCTCCATATACGTAATC 238
QY 217 GGAATAGACATATGCTGTAACCTTCTCAACTATGCTTCCTGCTGCTGCTGCTG 276
DB 239 GGAATAGACATATGCTGTAACCTTCTCAACTATGCTTCCTGCTGCTGCTGCTG 298
QY 277 TTTTCAGTCACTCAACGAGAACTATAAAGGTGCTTCGTGCTCCATATACGTAATC 336
DB 299 TTTTCAGTCACTCAACGAGAACTATAAAGGTGCTTCGTGCTCCATATACGTAATC 358
QY 337 CCTTCCTTCAACACCTTGCATCGGTTTACTGTCAGTCACTGCTGCTGCTGCTGCT 396
DB 359 CCTTCCTTCAACACCTTGCATCGGTTTACTGTCAGTCACTGCTGCTGCTGCTGCT 418
QY 397 GCTTATTTGCTGCAAGTGGTGTCTTCACTTACCAGTCTCTACTTCTCCCTCATCTG 456
DB 419 GCTTATTTGCTGCAAGTGGTGTCTTCACTTACCAGTCTCTACTTCTCCCTCATCTG 478
QY 457 GCTTCACAGTGGCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT 516
DB 479 GCTTCACAGTGGCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT 538
QY 517 CCTTTGATAGTCAGTCTTGTGCTTCTCCTCACTGATCTCTGCTTCTTGTGGTCAAC 576
DB 539 CCTTTGATAGTCAGTCTTGTGCTTCTCCTTACGGTTTCTCTGCTTCTCTGCTGCT 598
QY 577 ACTGATTGAGAACTCAACTAATGATGATGATGATGATGATGATGATGATGATGAT 636
DB 599 ACTGATTGAGAACTCAACAAATGATGATGATGATGATGATGATGATGATGATGAT 658

QY 637 ACCATCGTGTCTTCGCTGGGATTGGACTGTCTACTATCTCTGATACAAATGCTCTTCAGG 696
DB 659 ACCATCGTGTCTTCGCTGGGATTGGACTGTCTACTATCTCTGATACAAATGCTCTTCAGG 718
QY 697 AAAGTTTTCACGAAG 711
DB 719 AAAGTTTTCACGAAG 733
RESULT 11
BX664674/c 721 bp DNA linear STS 14-OCT-2003
LOCUS Arabidopsis thaliana transposon insertion STS AT_5.3478, sequence
DEFINITION tagged site.
ACCESSION BX664674
VERSION BX664674.1 GI:37665166
KEYWORDS STS; STS, sequence tagged site.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
1
REFERENCE Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B., Walsh,S.,
AUTHORS Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 721)
AUTHORS Clarke,J.H.
TITLE Direct Submission
JOURNAL Submitted (14-OCT-2003) Clarke J.H., John Innes Centre, Colney
Lane, Norwich, NR4 7UJ, UK
COMMENT AT denotes an activation tag dissociation transposon within a
single line. ET an enhancer trap dissociation transposon, GT a gene
trap dissociation transposon, MT a mis-expression enhancer trap
dissociation transposon, SM a defective suppressor mutator
transposon. 3 denotes a sequence derived from the 3' end of the
transposon, 5 denotes a sequence derived from the 5' end of the
transposon, BSRG GARNET, ATIS project
On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
code: N101051.
Location/Qualifiers
1..721
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/var_tely="Landsberg erecta NASC stock code NW20"
/cultivar="Thil3xTn25"
/db_xref="taxon:3702"
/clone="AL021713"
1..721
/standard_name="AT_5.3478"
ST
ORIGIN
Query Match 45.7%; Score 546.2; DB 11; Length 721;
Best Local Similarity 99.5%; Pred. No. 1.1e-141;
Matches 548; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 644 GTGCTTCGCTGGGATTGGACTGTCTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTT 703
DB 707 GGGCTTCGCTGGGATTGGACTGTCTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTT 648
QY 704 TCAGGAGCATACATCTCAGCAGTCCAGGACTTGGCCATTTACCAGTCTCTAGTTGGGA 763
DB 647 TCAGGAGCATACATCTCAGCAGTCCAGGACTTGGCCATTTACCAGTCTCTAGTTGGGA 588
QY 764 GTTGTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGAGTGGGAAACTTTTGCAAGTGAGA 823
DB 587 GTTGTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGAGTGGGAAACTTTTGCAAGTGAGA 528
QY 824 TCAGGAGCATACAACTCGGAAAGTGTATATGTTTACCTTTAGCTTCGCGCAGCTATTT 883
DB 527 TCAGGAGCATACAACTCGGAAAGTGTATATGTTTACCTTTAGCTTCGCGCAGCTATTT 468

Qy 1157 ACACAA 1162
|
Db 1043 AGTCAA 1048

Search completed: November 1, 2004, 15:44:43
Job time : 5297.13 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:40:28 ; Search time 678.972 Seconds
(without alignments)
9996.736 Million cell updates/sec

Title: US-09-913-767-4
Perfect score: 1293
Sequence: 1 ggaagtcctcttagtggt.....tatgatcaaacatatattcc 1293

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1293	100.0	1293	3	Aaa97922 A. thalia
2	1152	89.1	3387	6	Abz14407 Arabidops
3	1152	89.1	3387	8	Adm68246 Arabidops
4	816.6	63.2	1194	3	Aaa97923 A. thalia
5	435.2	33.7	1071	3	Aaa97925 A. thalia
6	410	31.7	1047	6	Adg88190 A. thalia
7	410	31.7	1081	3	Aaa97924 A. thalia
8	354.4	27.4	592	10	Abx56692 Arabidops
9	294.4	22.8	1145	3	Aaa97921 A. thalia
10	294.4	22.8	1292	3	Aac51119 Arabidops
11	294.4	22.8	1295	3	Aac37289 Arabidops
12	289	22.4	2660	3	Aac47974 Arabidops
13	246.6	19.1	8398	6	Abn85767 Arabidops
14	230.8	17.8	1403	3	Aac37652 Arabidops
15	227.2	17.6	2175	8	Ada70748 Rice gene
16	214.4	16.6	1152	8	Ada70213 Rice gene
17	211.6	16.4	492	3	Aac36845 Arabidops
18	157	12.1	1049	3	Aaa97920 A. thalia
19	140.6	10.9	1155	8	Adm69686 Rice gene
20	128	9.9	386	6	Abq85215 Arabidops
21	126.8	9.8	1225	3	Aaa97919 A. thalia

22	126	9.7	2586	3	AAC44184	Aac44184 Arabidops
23	116	9.0	1040	8	ADA69568	Ada69568 Rice gene
24	115.2	8.9	332	12	ADP92185	Adp92185 Cotton ex
25	89.4	6.9	5520	5	AAC68924	Aac68924 FUR1 rece
26	69.2	5.4	1173	12	ADN72520	Adn72520 Thale cre
27	67.8	5.2	774	6	ABN99127	Abn99127 Arabidops
28	67	5.2	477	3	AAC37273	Aac37273 Arabidops
29	62.6	4.8	1200	8	ADA70627	Ada70627 Rice gene
30	51.8	4.0	493	3	AAC36831	Aac36831 Arabidops
31	49.2	3.8	1418	3	AAC46395	Aac46395 Arabidops
32	47.6	3.7	1421	3	AAC39779	Aac39779 Arabidops
33	46.6	3.6	446	3	AAC37188	Aac37188 Arabidops
34	46	3.6	2613	10	ADG32560	Adg32560 Generic m
35	45.6	3.5	174424	6	ABL68122	Ab168122 Ovary can
36	45.6	3.5	181343	12	ADQ19573	Adq19573 Human sof
37	45.2	3.5	2000	8	ADA71938	Ada71938 Rice gene
38	43.6	3.4	2613	10	ADG32563	Adg32563 Generic h
39	41.6	3.2	17934	6	ABL33718	Ab133718 Human inm
40	41.4	3.2	300	2	AAZ14317	Aaz14317 Human gen
41	41.2	3.2	97835	6	ABK84796	Abk84796 Human CDR
42	41	3.2	31670	8	AAD56093	Aad56093 Human CCR
43	41	3.2	31670	9	ADA02455	Ada02455 Human CCR
44	41	3.2	31670	10	ADB72194	Adb72194 Human CCR
45	40.6	3.1	543	12	ACH70417	Ach70417 Human gen

ALIGNMENTS

RESULT 1
AAA97922
ID AAA97922 standard; DNA; 1293 BP.
AC AAA97922;
XX
DT 19-JAN-2001 (first entry)
XX
DE A. thaliana PUP1 DNA #4.
XX
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.
XX
OS Arabidopsis thaliana.
XX
PN DE19907209-A1.
XX
PD 24-AUG-2000.
XX
PF 19-FEB-1999; 99DE-01007209.
XX
PR 19-FEB-1999; 99DE-01007209.
XX
(FROM/) FROMMER W.
XX
GI Gillissen B, Buerkle L, Andre B, Frommer WB;
PI WPI; 2000-566202/53.
DR
PT Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transporter, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.
PS
PS Claim 1f; Page 13; 24pp; German.
XX
CC This invention describes a novel nucleic acid encoding a plant nucleobase
CC transporter (I). (I) is produced by complementation of a nucleobase
CC transport (NBT)-defective host cell with a plant gene bank by selection
CC of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
CC expression of (II) (when in antisense orientation), and to produce
CC transgenic crop plants. The transgenic plants have modified nucleobase
CC transport properties, e.g. altered affinity and substrate specificity

CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
xx
SQ Sequence 1293 BP; 345 A; 288 C; 247 G; 413 T; 0 U; 0 Other;

901	901	AAA	AACTTTAAC	AGT	GAG	ATGG	AAAC	ACT	CAAA	ACTGGG	GAAG	TGCC	ATAC	GTAT	TAC	GT	960
961	961	TTG	GCCTCG	ATAG	TACT	ATTTCT	CGCA	AGTCT	TAC	ACCA	ATGG	CGGTCG	TGG	CACTG	ATCTTT		1020
961	961	TTG	GCCTCG	ATAG	TACT	ATTTCT	CGCA	AGTCT	TAC	ACCA	ATGG	CGGTCG	TGG	CACTG	ATCTTT		1020
1021	1021	GAG	TCATCTT	CTG	TTTCT	CCAA	TTCC	AA	TTCC	AA	TTCC	AA	TTCC	AA	TTCC		1080
1021	1021	GAG	TCATCTT	CTG	TTTCT	CCAA	TTCC	AA	TTCC	AA	TTCC	AA	TTCC	AA	TTCC		1080
1081	1081	GTAG	CAGT	GAT	TGTTT	TCAT	GAT	ATAA	ATG	AA	CGCGT	CAAA	AGATCT	TCT	CC	ATCAT	1140
1081	1081	GTAG	CAGT	GAT	TGTTT	TCAT	GAT	ATAA	ATG	AA	CGCGT	CAAA	AGATCT	TCT	CC	ATCAT	1140
1141	1141	GCT	ATCTG	GGG	GATTC	ATT	TCT	AT	CAG	CAC	TAC	CTCG	ACG	AAAA	AGAG	TTC	1200
1141	1141	GCT	ATCTG	GGG	GATTC	ATT	TCT	AT	CAG	CAC	TAC	CTCG	ACG	AAAA	AGAG	TTC	1200
1201	1201	ACT	AGCC	AC	AC	AA	GT	CTCT	G	TAG	GAG	ATCC	T	CAT	C	TAC	1260
1201	1201	ACT	AGCC	AC	AC	AA	GT	CTCT	G	TAG	GAG	ATCC	T	CAT	C	TAC	1260
1261	1261	AAC	AT	CAT	AT	AGT	GT	AT	GAT	CAAA	AC	AT	TT	TC			1293
1261	1261	AAC	AT	CAT	AT	AGT	GT	AT	GAT	CAAA	AC	AT	TT	TC			1293

RESULT 2	
ABZ14407	
ID	ABZ14407 standard; DNA; 3387 BP.
XX	
AC	ABZ14407;
XX	
DT	21-JAN-2003 (first entry)
XX	
DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 2212.
XX	
KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX	
OS	Arabidopsis thaliana.
XX	
PN	WO200216655-A2.
XX	
PD	28-FEB-2002.
XX	
PF	24-AUG-2001; 2001WO-US026685.
XX	
PR	24-AUG-2000; 2000US-0227866P.
PR	26-JAN-2001; 2001US-0264647P.
PR	22-JUN-2001; 2001US-0300111P.
XX	
PA	(SCRI) SCRIPPS RES INST.
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.
XX	
PI	Harper JF, Kreps J, Wang X, Zhu T;
XX	
DR	WPI : 2002-304127/34.

CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office

SQ Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;

```
QY 22 ATGGAATAAATCAAGTAATCTATGTCTCAATGGTAAGCAAGATGCATCTCGTAGTGTAGAT 81
Db 1 ATGGAATAAATCAAGTAATCTATGTCTCAATGGTAAGCAAGATGCATCTCGTAGTGTAGAT 60
QY 82 TACTGTGATCTCTTTTCGCTAACTGTGTGTTTGTGATGTTTTTCAGATCATATAACAACATAGAA 141
Db 61 TACTGTGATCTCTTTTCGCTAACTGTGTGTTTGTGATGTTTTTCAGATCATATAACAACATAGAA 120
QY 142 GCAAACTTAAACAGGTTCAGGAGAAATGAATACCAACATGGAATTCGAATCTTCGTCGGTA 201
Db 121 GCAAACTTAAACAGGTTCAGGAGAAATGAATACCAACATGGAATTCGAATCTTCGTCGGTA 180
QY 202 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTTCCATTTACGTGTTCTTTGTCTCTT 261
Db 181 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTTCCATTTACGT----- 227
QY 262 GCTTGCCCAAGCACTTTCTACAAATTTTGGGCAGAGTTTACTATGAAATGGTGGGAAGAGT 321
Db 228 -----AGTTTACTATGAATGGTGGGAAGAGT 255
QY 322 ACATGGATGGAAACATTTGTCCAACTAATCGGCTTCCCTGTCTGTCTCTTCCTTCGCTTC 381
Db 256 ACATGGATGGAAACATTTGTCCAACTAATCGGCTTCCCTGTCTGTCTCTTCCTTCGCTTC 315
QY 382 TTTTCCCAACCAAAATCCCAACCAACAGAGAGATTTTCAGAAAGTTCTCTTCCTTC 441
Db 316 TTTTCCCAACCAAAATCCCAACCAACAGAGAGATTTTCAGAAAGTTCTCTTCCTTC 375
QY 442 ACCATTCTTGATCAGTTTACATCGTTACTCGGACTATTAGTGTCTGCTAACTCTTATATG 501
Db 376 ACCATTCTTGATCAGTTTACATCGTTACTCGGACTATTAGTGTCTGCTAACTCTTATATG 435
QY 502 TCCTCTGTGGTTTACTATACATACCAAGTTTCTACTTTCTCCCTCATCTTGGCCTCACAA 561
Db 436 TCCTCTGTGGTTTACTATACATACCAAGTTTCTACTTTCTCCCTCATCTTGGCCTCACAA 495
QY 562 TTGGCCTTCACGCTTCTCTCATATTTTCTAACTCGCAGAGAGTTTCACACCTTTTCATT 621
Db 496 TTGGCCTTCACGCTTCTCTCATATTTTCTAACTCGCAGAGAGTTTCACACCTTTTCATT 555
QY 622 GTGAATCTCTGTCTTCTCTTACTATTTCTCTGCTGCTCCCTCGTGGTCAACACATGATTCG 681
Db 556 GTGAATCTCTGTCTTCTCTTACTATTTCTCTGCTGCTCCCTCGTGGTCAACACATGATTCG 615
QY 682 GAAACACAGCAAAAGTGTCTAGAGTAAATATATGTGATAGGGATAAATGTACCATTTGGT 741
Db 616 GAAACACAGCAAAAGTGTCTAGAGTAAATATATGTGATAGGGATAAATGTACCATTTGGT 675
QY 742 GCTTCTGCTGGATGGATTCGCTGCTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
Db 676 GCTTCTGCTGGATGGATTCGCTGCTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735
QY 802 AAGAGCAAAACATTTCTCAACGCTCACTGACTTGGTGGCTTACCAATCTCTAGTTGCAAGC 861
Db 736 AAGAGCAAAACATTTCTCAACGCTCACTGACTTGGTGGCTTACCAATCTCTAGTTGCAAGC 795
QY 862 TGTGTGTTCTCATAGACTTTTCGAAGCGGGAGTGGAAACCTTTTAAACAGTGAGATG 921
Db 796 TGTGTGTTCTCATAGACTTTTCGAAGCGGGAGTGGAAACCTTTTAAACAGTGAGATG 855
QY 922 GAAACTACAACCTGGGGAAGTGCATACGTTATGACTTTGGCTCGATAGCTATTTCC 981
Db 856 GAAACTACAACCTGGGGAAGTGCATACGTTATGACTTTGGCTCGATAGCTATTTCC 915
QY 982 TGGCAAGTCTACACATTTGGGCTCGTGGGACTGATCTTTGAGTCACTCTCTGTGTTCTCC 1041
Db 916 TGGCAAGTCTACACATTTGGGCTCGTGGGACTGATCTTTGAGTCACTCTCTGTGTTCTCC 975
QY 1042 AATTCCTAACTGCTGTGGGATTCGCTATAGTTCCAGTTGTAGGAGTGAATTTTTCAT 1101
Db 976 AATTCCTAACTGCTGTGGGATTCGCTATAGTTCCAGTTGTAGGAGTGAATTTTTCAT 1035
```

```
QY 1102 GATAAAATGAACGGTCAAAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCATTTC 1161
Db 1036 GATAAAATGAACGGTCAAAAGATCTTCTCCATCATTTTAGCTATCTGGGATTCATTTC 1095
QY 1162 TTTGTCTATCAGCACTACCTCGACGAAAGAGTTGAAGACTAGCCACACAAGTCTCTGTA 1221
Db 1096 TTTGTCTATCAGCACTACCTCGACGAAAGAGTTGAAGACTAGCCACACAAGTCTCTGTA 1155
QY 1222 GGAGATCCTCATCTACTACTCTGCTGAGGAAGGTCACAAACATACATAGTG 1273
Db 1156 GGAGATCCTCATCTACTACTCTGCTGAGGAAGGTCACAAACATACATAGTG 1207

RESULT 4
AAA97923
ID AAA97923 standard; DNA; 1194 BP.
AC AAA97923;
XX
DT 19-JAN-2001 (first entry)
XX
DE A. thaliana PUP1 DNA #5.
XX
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;
flowering behaviour; senescence; pesticide distribution; ds.
XX
OS Arabidopsis thaliana.
XX
FN DE19907209-A1.
XX
PD 24-AUG-2000.
XX
PF 19-FEB-1999; 99DB-01007209.
XX
PR 19-FEB-1999; 99DB-01007209.
XX
PA (FROM/) FROMMER W.
XX
PI Gillissen B, Buerkle L, Andre B, Frommer WB;
XX
WP1; 2000-566202/53.
XX
DR
CC This invention describes a novel nucleic acid encoding a plant nucleobase
CC transporter (I). (I) is produced by complementation of a nucleobase
CC transporter (NBT)-defective host cell with a plant gene bank by selection
CC of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (I) in prokaryotic or eukaryotic cells, for inhibiting
CC expression of (II) (when in antisense orientation), and to produce
CC transgenic crop plants. The transgenic plants have modified nucleobase
CC transport properties, e.g. altered affinity and substrate specificity
CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX
SQ Sequence 1194 BP; 298 A; 279 C; 231 G; 386 T; 0 U; 0 Other;

Query Match 63.2%; Score 816.6; DB 3; Length 1194;
Best Local Similarity 81.8%; Pred. No. 6.1e-233;
Matches 971; Conservative 0; Mismatches 204; Indels 12; Gaps 2;

QY 107 TGTGTTTGTGTTTTCAGATCATACATAGCAACCAACCTTAACAGGTTCAGGAGGAA 166
Db 20 TGAAGTGTAAATTTTTCAGGTGACCAAGAACTTAGAAGCAAAACCTTATAGATCATGAGGTGG 79
QY 167 TGAATACCACCATGGAAATCGAATCTTTCGTCGCTACCTCAATCGAAGAACTATAAGAAAT 226
```

Db 80 TAACT-----GAATCATCATCATCAGCTGTGCTCAACCGAGAACTATAAAGGT 130
Qy 227 GGCTTCGTATTTCCATTTAGTGTCTTTCTCTCTGCTGCGCAAGCACTTTCTACAAATT 286
Db 131 GGCTTGGTCTCCATATAGTAACTTTTGTCTCTTTTGGCCAGCACTAGCTACAAATC 190
Qy 287 TGGGAGAGTTTACTATGAAATGGTGGGAAGTACATGGATGGGAACACTTTGTCCAAC 346
Db 191 TGGGTAGATTGTACTATGAAATGGAGAAATAGCACATATGTGGTAACACTTTCTCAAC 250
Qy 347 TAATCGGCTCCCTGTTCTGTTTCTCTCGCTTCTTTTCCCAACCAAAATCCCAAC 406
Db 251 TCATTGGCTTCCCTGACTGTTCTGTTCCGCTTCTTTCTCGAATCAGCAACCAAAAT 310
Qy 407 CAACAGAAGCAGATTTTCAGAAAGTCTCTCTCCCTTCAACCATTTCTGATCAGTTTACATCG 466
Db 311 CAACAGATACAAATTTTCAGTCAGTCCCTTCCCTTCAACCCCTTGCATCGTTTACTGT 370
Qy 467 TTACTGGACTATTAGTGTCTGCTAACTCTTATATGTCTCTGTTGTTTACTATATTAC 526
Db 371 GCACCTGGACTGTAGTGTGCGCTTATGCTTATTTGTCTGCAGTAGGTTGCTCTACTAC 430
Qy 527 CAGTTTCTACTTTCTCGCTCATCTTGGCCCTCACAATTTGGCTTCACTGCTTCTTCTCAT 586
Db 431 CAGTCTCTACTTTCTCCCTCATCTTGGCCCTCAAGTTGGCTTCACTGCTTCTTCTCAT 490
Qy 587 ATTTTCTAAACTCGCAGAGTTTCACACCTTTTCAATTCGTGAATCTCTGTTTCTCTTACTA 646
Db 491 ATTTCTTAACCTCGCAAAAGTTTCACTCCCTTGTGATGTCAGTTCTTTGCTTCTCTCACTG 550
Qy 647 TTTCTCTGCGCTTCTCGTGGTCAACACTGATTCGGAACACACAGCAAAAGTGTCTAGAG 706
Db 551 TATCTCTGCTCTTCTTGTGTTCAACACTGATTCAGAAACTCAACTATGATATCTAGAG 610
Qy 707 TAAATATGTATAGGATTAATGTACCAATGGTGTCTCTGCTGGGATGGATGCTGCTG 766
Db 611 TACAGTATGTGATCGGGTTTATATGATCCATCGTGTCTTCCGCTGGGATGGACTGTAC 670
Qy 767 TATCCCTGTGTACAACTGATCCTCAGGAAGTTTAAAGAAAGCAAACTCTCAACGGTCA 826
Db 671 TATCTGTATACAAATGCTCTCAGGAAGTTTTCAGGAAGCATATCTCAGCAGTCA 730
Qy 827 CTGACTTGGTTCGCTTACCACTCTAGTTGCAAGTGTGTGGTCTCTCATAGGACTTTTCG 886
Db 731 CGGACTTGGCAATTTACCAGTCTCTAGTTGGGAGTTGTGTAGTTCTCATAGGACTTTTG 790
Qy 887 CAAGCGGGAGTGGAAACTTTTAAAGTGAGATGGAAACTTACAAACTGGGGAAGTGC 946
Db 791 CAAGTGGAGTGGGAACTTTTCCCAAGTGAGATGAGAACTTACAACTCGGGAAGTGT 850
Qy 947 CATACGTTATGACTTTGGGCTCGATAGCTATTTTCTGGCAAGTCTACACCATTTGGGCTG 1006
Db 851 CATATGTTTGTAGCTTGGGCTCGGAGCTATTTCTCGCAAGTCTACACTTTGGTCTTG 910
Qy 1007 TGGGACTGATCTTTGAGTCACTCTTGTGTCTTCTCCAAATTCCTAACTGTGTGGGATTC 1066
Db 911 TGGGATTTGATCTTCGAGTCACTCTCTGTGTCTTCCAAATTCCTAACTGATGAGTTCG 970
Qy 1067 CTATAGTTCCAGTTGTAGAGTGAATTTTTCATGATGAATGAATGAATGAATGAATGAAT 1126
Db 971 CTATAGTTCCAGTTGTGGGAGTGAATTTTTCATGATGAATGAATGAATGAATGAATGAAT 1030
Qy 1127 TCTCCATCTTTTGTAGTATCTGGGATTCATTTTCTCTATCTATCAGCACTACCTCGAG 1186
Db 1031 TCTCCATTTTGTAGTATCTGGGCTTCTTTTCTATCTCTATCAGCACTACCTCGAG 1090
Qy 1187 AAAAGAAGTTGAAGTACGACCAACAAGTCTGTAGGATCTCTATCTACTACTGCTG 1246
Db 1091 AAAAGAAGTTGAATCTAGCCACACAAGTGTGTAGGATCTTCTATCTTCTGTTG 1147
Qy 1247 AGAAGGTCAACAAACATACATAGTGTATGATCAAAACATATTTCC 1293

Db 1148 AGGAAGGTCAACAAACATACAAAGTGTGTGATCAAGCATATTTCC 1194
RESULT 5
AAA97925
ID AAA97925 standard; DNA; 1071 BP.
XX AC AAA97925;
XX DT 19-JAN-2001 (first entry)
XX DE A. thaliana PUP1 DNA #7.
XX KW PUP1; transgenic plant; nucleobase transporter; apical dominance;
XX KW flowering behaviour; senescence; pesticide distribution; ds.
XX OS Arabidopsis thaliana.
XX PN DE19907209-A1.
XX PD 24-AUG-2000.
XX PF 19-FEB-1999; 99DE-01007209.
XX PR 19-FEB-1999; 99DE-01007209.
XX PA (FROM/) FROMMER W.
XX PI Gillissen B, Buerkle L, Andre B, Frommer WB;
XX WPI; 2000-566202/53.
XX Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transport, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.
XX PS Claim 2; Page 15; 24pp; German.
XX CC This invention describes a novel nucleic acid encoding a plant nucleobase
CC transporter (I). (I) is produced by complementation of a nucleobase
CC transport (NBT)-defective host cell with a plant gene bank by selection
CC of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
CC expression of (II) (when in antisense orientation), and to produce
CC transgenic crop plants. The transgenic plants have modified nucleobase
CC transport properties, e.g. altered affinity and substrate specificity
CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX Sequence 1071 BP; 307 A; 236 C; 208 G; 320 T; 0 U; 0 Other;
SQ
Query Match 33.7%; Score 435.2; DB 3; Length 1071;
Best Local Similarity 65.8%; Pred. No. 5.2e-119;
Matches 651; Conservative 0; Mismatches 333; Indels 6; Gaps 1;
Qy 264 TTGCCAAGCAGCTTCTCAAAATTTGGGAGAGTTCATGAAATGGTGGGAAGATAC 323
Db 51 TGGCCAAATCAGTTGTACAAATTTCTGGGAGACTATATGAAATGGAGGAACAGCAA 110
Qy 324 ATGGATGGGAACACTTGTCCCAACTAATCGGCTTCCCTGTCTGTTTCTTCCGCTCTT 383
Db 111 ATGGCTAGCAAGGTAGTTTTCAGCTTGTAGGCTTCTTCTTCTTCCATATCATCTCTT 170
Qy 384 TTCCCAACCAAAATCCCAACCAACCAAGAGCAGATTTTCAGAAAGTCTCTTCTTCCAC 443
Db 171 GTCAGTCAAAACACATACAACTCAGAGATG-----GCAATTAACCTCACTTAG 224
Qy 444 CATTCTTGGATCAGTTTACATCGTTTACTTGGAGTATAGTGTCTGCTAACTCTTATATGTC 503
Db 225 GAACCGTGCATTAGTTTACATAGTCTTGGAGTCTTCTTGTAGGAGCAGCTTGTCTACCTATA 284

Db 743 TTAGAACCTAGTGTGACAGCTGTACCTGGCAGGTATCTCCATCGGTGGCACAGAC 802
Qy 1013 TGAATCTTTGAGTCATCTTCTGTGTCTCCAAATCCATAACTGCTGTGGATTGCTATAG 1072
Db 803 TGAATCTTGAGCTCTCCCTCTCTATCTCAATGCAATGAAGGTTTGGGACTCCAGTGG 862
Qy 1073 TTCAGTTGTAGCAGTGAATGTTTTCATGATATAAATGAAGCGTCAAGATCTTCTCCA 1132
Db 863 TTCATATCTTGGCTGTAAATCAATTTTCCATGACAAAATGAATGGGTAAAGGTGATTTCTA 922
Qy 1133 TCATTTAGCTATCTGGGATTCATTTTCTATCTATCTACACTACTCTGCACAAAGA 1192
Db 923 TGAATCTAGCTATTTGGGTTTCACTTCTATGCTACCAACATATCTTGATGACAAA 982
Qy 1193 AGTTGAAGACTAGCCACAAAGTCTGTGAGAGATCCCTCATCTACTACTGCTGAGGAAG 1252
Db 983 ACTTGAAGAAAAATCATGAATCAACAAACAGAAATCCCTGACCCACAGAGAGAAG 1042

RESULT 8

ABX56692
ID ABX56692 standard; DNA; 592 BP.
XX AC ABX56692;
XX DT 20-FEB-2003 (first entry)
XX DE Arabidopsis thaliana polynucleotide #44.
XX KW Thale cress; gene; ds; genetic manipulation; plant; biosynthesis;
KW genetic modification; environmental stress; disease resistance;
KW fungicide; insecticide; stress tolerance.
XX OS Arabidopsis thaliana.
XX PN US2002040489-A1.
XX PD 04-APR-2002.
XX PF 26-JAN-2001; 2001US-00770152.
XX PR 27-JAN-2000; 2000US-0178503P.
XX PA (GORL/) GORLACH J.
PA (ANY/) AN Y.
PA (HAMI/) HAMILTON C M.
PA (PRIC/) PRICE J L.
PA (RAIN/) RAINES T M.
PA (YUY/) YU Y.
PA (RAME/) RAMEAKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRIC/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX Gortlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX WPI; 2003-110410/10.
XX Novel Arabidopsis thaliana nucleic acid useful for identifying homologous
PT or related genes, and to create genetically modified and transgenic

PT organisms, such as plant cells and plants.
XX Claim 1; SEQ ID NO 44; 45pp; English.
XX The invention relates to Arabidopsis thaliana nucleic acid sequences. The
CC DNA sequences and the polypeptides they encode are useful for identifying
CC homologous or related genes, for producing compositions that modulate the
CC expression or function of the polypeptides, for mapping functional
CC regions of the protein, in diagnosis, for studying associated
CC physiological pathways, for genetic manipulation of cells, preferably
CC plant cells, in screening assays of various plant strains to determine
CC the strains that are capable of withstanding a particular disease or
CC environmental stress, for enhancing or inhibiting production of
CC biosynthetic products in plants and to create genetically modified and
CC transgenic organisms, such as plant cells and plants. Transgenic plants
CC are useful for introducing or improving disease resistance and stress
CC tolerance in plants, screening biologically active agents, such as
CC fungicides and insecticides, and for elucidating biochemical pathways.
CC Sequences ABX56649-ABX57647 represent Arabidopsis thaliana
CC polynucleotides of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html
XX SQ Sequence 592 BP; 135 A; 158 C; 96 G; 203 T; 0 U; 0 Other;
Query Match 27.4%; Score 354.4; DB 10; Length 592;
Best Local Similarity 81.8%; Pred. No. 5.3e-95;
Matches 409; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
Qy 180 GGAAATCGAATCTTCGTCCGTACTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTC 239
Db 93 GGAACCTGAATCATTTTCAGTACTCTCAACGAGAACTGTAGAGGTGGCTCCGTGCTC 152
Qy 240 CATTACGTGTTCTTTGTCCTTGGCTTGCAGCACTTTCTACAAATTTTGGGCGAGTTTA 299
Db 153 CATATACGAATCTTTGTCTATCTTCTGCCAACCACTTGTCTACAGTTCTGGGTAGACTGA 212
Qy 300 CTATGAAATGGTGGGAGAGTACATGGATGGGAACTGTCCAACTAATCGGCTTCCC 359
Db 213 CTATGAAATGGAGGGGAAAGCACAATATGGTAAACACTTCTTCACTCATTTGGCTTCCC 272
Qy 360 TGTTCGTCTCTCTCCGCTTCTTTTCCAAACCAAAATCCCAACCAACAGACAGA 419
Db 273 TGTACTGATTTCTCTTCGCTTCTTTCTCGAATCAGGCAACCCAAATCAACAGATACAA 332
Qy 420 TTTCAGAAAGTTCTCTTTCCTTCCACTTCTTGATCAGTTTACATCGTTACTGGACTATT 479
Db 333 TTTCAGTCACTCCCTTCTCTTCCACCCCTTGCATCGGTTTACTTGTGCATCGGACTGCT 392
Qy 480 AGTCTCTGCTAACTCTTATATGCTCTGTTGGTTTACTACTTACCAGTTTCTACTTTT 539
Db 393 AGTCTGCTTATGCTTTATTTGTCGAGTTGGGTGCTTTACTTACAGTCTCTACTTTT 452
Qy 540 CTCCTCATCTTGGCCCTCACAATTTGGCTTTCACCTGCTTCTCTCATATTTCTTAAACTC 599
Db 453 CTCCTCATCTTGGCCCTCAGATTGGCTTTCACCTGCTTCTCTCATATTTCTTAAACTC 512
Qy 600 GCAGAAAGTTCAACCTTTCAATTTGGAATTTCTGTTTCTCTTACTATTTCTCTGCGCT 659
Db 513 GCAAAAGTTCACTCTTGTATGATGATCAATTTCTTGTCTCTCTACGGTTTCTCTGCGCT 572
Qy 660 CCTCGTGGTCAACACTGATT 679
Db 573 CCTCGTGGTCAACACTGATT 592
RESULT 9
AAA97921
ID AAA97921 standard; DNA; 1145 BP.
XX AC AAA97921;
XX DT 19-JAN-2001 (first entry)

PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131443P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134258P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136382P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139432P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140333P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 26-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.

PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161933P.	
PR	28-OCT-1999;	99US-0162142P.	
Query Match 22.8%; Score 294.4; DB 3; Length 1292;			
Best Local Similarity 55.6%; Pred. No. 6.6e-77;			
Matches 593; Conservative 0; Mismatches 461; Indels 12; Gaps 1;			
QY	141	AGCAACCTAACAGGTCAGAGGAAATGAATACACCATGGGAAATCGAATCTTCGTCCT	200
DB	39	AGAACTCAAGAACTCCATCTTCACGTAATGCGAAACCTGGAAGGAAATTTTCAACAGA	98
QY	201	ACCTCAATCGAAGCATATAAGAAATGGCTTCCTATTTCATTTAGCTTCTTTCTCT	260
DB	99	AGAGAGAGTCACAGTACTCTTGGAGGTTAAGAGTGCTCTCTATGCACTCTCTCTT	158
QY	261	TGCTTCCAGCACTTTTACAAATTTTGGCGAGTTTACTATGAAATGGTGGGAAGAG	320
DB	159	AGCTGGAGACATAGCCACTCTCTAGTAGACTTTACTAGAAAGGCGGTAAAG	218
QY	321	TACATGGATGGGAACACTTGTCCAACTAATCGGCTCCCTGTTCTGTTCTCTCTCCGTT	380
DB	219	CACATGGCTCGAAACCTTGGTTTCAGCTTGTAGGTTTCCCTTAAACCTTCTTGTATTA	278
QY	381	CTTTTCCCAACCAAAATCCCAACCAACAGAGCAGATTTACAGAAAGTTCTCTCTT	440
DB	279	TTACTTAAAGCTTGAGCGGTCCAGACTAAACCAATTAACCAAAATCTCTCTCTT	338
QY	441	CACCATTTTGGATCAGTTTACATCTGTTACTGACTATTAGTCTGTCAACTCTTATAT	500
DB	339	CTTGACACTATCTTTAGTGTATATGGACTTGGCTGCTTGTGGACATGTATTTT	398
QY	501	GTCTCTGTGGTTTACTATACATTAACAGTTTCTACTTCTCTCTCTCTCTCTCTCT	560
DB	399	GTACTCATTTGGGCTACTTTACCTTCTGTCTCAACTTTCTTCTTCTCTCTCTCT	458
QY	561	ATTGGCTTCACTGCTCTTCTCTCATATTTCTTAAACTGCGAGAAGTTTCAACCTTT	620
DB	459	ATTGGCTTTTAAAGCGCTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTT	518
QY	621	TGTGAATTTCTGTTTCTCTTACTATTCTCTGCTGCTCTCTCTCTCTCTCTCTCT	680
DB	519	ACTCAATTCATTTGTTCTCTTAACCATATCTCTACACTTCTTGTATCAACATGAAC	578
QY	681	GGAA-----AACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGGATAAT	728

DB	579	AGAACTCTCCTCTTCTTACTTCAAGTCCGACCAAGTCAAGTATGTGATTGATACAT	638
QY	729	ATGTACCATTTGCTGCTGCTGGGATTGGATTGCTGCTATCCCTGGTACAACTGATCCT	788
DB	639	CTGCGCGGTGGTAGCTCAGCTGGTTATTTCTCTGGTGGTTCTTTTAAAGATTACCGGT	698
QY	789	CAGGAAGGTTTTAAAGAAACAAATCTCTCAACGGTCACTGACTTGGTGGCTTACCAATC	848
DB	699	CGAAAAGATTCTAAAGAAATACACATTCAAGGCTATTTTAGACATGCGCCACATATCCGTC	758
QY	849	TCTAGTTGCAAGCTGTGGTCTCATAGGACTTTTTCGCAAGCGGGAGTGGAACCTTT	908
DB	759	TATGGTAGTACTTGTGTAGTTTGTGTAGGACTTTTTCGAAAGTGGTGGGTAAGACT	818
QY	909	AACAAGTGAGATGGAACAACTACAACTGGGAAAGTGCATACGTTTATGACTTTGGCCTC	968
DB	819	GAGTACAGAAATGGAAGAGTTTCAACTAGGAAAGAGCTCATACATTTTGTATTAACATCGG	878
QY	969	GATAGCTATTTCTGGCAAGTCTACACCAATTTGGCGTCTGGGAGTATCTTTGAGTCATC	1028
DB	879	TTCAACGATATCATGCAAGCTTTGTTGATTGGAAGTGTGGTTGATTATCGAAGTTTC	938
QY	1029	TTCTGTGTTCTCAATTCATTAACCTGCTGTGGGATTCCTATAGTTCAGTTGTAGCAGT	1088
DB	939	ATCGCTTTTTCCTCAATGTCATAAGCACTCTTTGTTTACCAGTTGTGCTCTGTTCTGCTGT	998
QY	1089	GATTGTTTTCCTCATATAAATGAACGCGTCAAGAGATCTTCTCCATCATTTTGTAGCTATCG	1148
DB	999	TGTCTTCTTCCGTGATGAGTGAAGTGAATCAAGTTGGTTCGAATGTTTGTGGCAATCTG	1058
QY	1149	GGGATTCATTTCTATTGTTCTATCAGCACTACCTCGACGAAAAAGAAG	1194
DB	1059	GGGATTTGTTCTTATGTTATCAGCATTTATGTCATGATAGAAG	1104
RESULT 11			
AAC37289			
ID	AAC37289 standard; DNA; 1295 BP.		
XX	AAC37289;		
AC	AAC37289;		
XX	17-OCT-2000 (first entry)		
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 16846.		
DE	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX	Arabidopsis thaliana.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-00301439.		
XX	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	16-APR-1999; 99US-0128714P.		
PR	18-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-013049P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		

PR 30-APR-1999;	99US-0132048P.	PR 21-JUL-1999;	99US-0145086P.
PR 30-APR-1999;	99US-0132407P.	PR 21-JUL-1999;	99US-0145088P.
PR 04-MAY-1999;	99US-0132484P.	PR 22-JUL-1999;	99US-0145085P.
PR 05-MAY-1999;	99US-0132485P.	PR 22-JUL-1999;	99US-0145087P.
PR 06-MAY-1999;	99US-0132486P.	PR 22-JUL-1999;	99US-0145089P.
PR 07-MAY-1999;	99US-0132487P.	PR 22-JUL-1999;	99US-0145192P.
PR 11-MAY-1999;	99US-0132863P.	PR 23-JUL-1999;	99US-0145145P.
PR 14-MAY-1999;	99US-0134256P.	PR 23-JUL-1999;	99US-0145218P.
PR 14-MAY-1999;	99US-0134218P.	PR 23-JUL-1999;	99US-0145224P.
PR 14-MAY-1999;	99US-0134219P.	PR 26-JUL-1999;	99US-0145276P.
PR 14-MAY-1999;	99US-0134221P.	PR 27-JUL-1999;	99US-0145913P.
PR 18-MAY-1999;	99US-0134370P.	PR 27-JUL-1999;	99US-0145918P.
PR 18-MAY-1999;	99US-0134768P.	PR 27-JUL-1999;	99US-0145919P.
PR 19-MAY-1999;	99US-0134941P.	PR 28-JUL-1999;	99US-0145951P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146386P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146388P.
PR 24-MAY-1999;	99US-0135622P.	PR 02-AUG-1999;	99US-0146389P.
PR 25-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147038P.
PR 27-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147204P.
PR 28-MAY-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147192P.
PR 03-JUN-1999;	99US-0137528P.	PR 05-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138884P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 24-JUN-1999;	99US-0140629P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140823P.	PR 07-SEP-1999;	99US-0152363P.
PR 29-JUN-1999;	99US-0140991P.	PR 10-SEP-1999;	99US-0153070P.
PR 30-JUN-1999;	99US-0141287P.	PR 13-SEP-1999;	99US-0153758P.
PR 01-JUL-1999;	99US-0141842P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0142154P.	PR 16-SEP-1999;	99US-0154039P.
PR 02-JUL-1999;	99US-0142055P.	PR 20-SEP-1999;	99US-0154779P.
PR 06-JUL-1999;	99US-0142390P.	PR 22-SEP-1999;	99US-0155139P.
PR 08-JUL-1999;	99US-0142803P.	PR 23-SEP-1999;	99US-0155486P.
PR 09-JUL-1999;	99US-0142920P.	PR 24-SEP-1999;	99US-0155659P.
PR 12-JUL-1999;	99US-0142977P.	PR 28-SEP-1999;	99US-0156458P.
PR 13-JUL-1999;	99US-0143542P.	PR 29-SEP-1999;	99US-0156596P.
PR 14-JUL-1999;	99US-0143624P.	PR 04-OCT-1999;	99US-0157117P.
PR 15-JUL-1999;	99US-0144005P.	PR 05-OCT-1999;	99US-0157753P.
PR 16-JUL-1999;	99US-0144085P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144086P.	PR 07-OCT-1999;	99US-0158029P.
PR 19-JUL-1999;	99US-0144325P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144331P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.

PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161933P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match			22.8%; Score 294.4; DB 3; Length 1295;
Best Local Similarity			55.6%; Pred. No. 6.6e-77;
Matches 593; Conservative			0; Mismatches 461; Indels 12; Gaps 1;
QY	141	AGCAAACTTAACAGGTGACGAGGAAATGAATACCACTGGAATCGAATCTTCGTCCGT	200
DB	40	AGAAACTCAAGAACTCCATCTTACGTAATGCGAACCTGAAGGGAATTTTCAACAGA	99
QY	201	ACCTCAATCGAAGACTATAAGAAATGGCTTCGTATTTCCATTTAGTGTTCTTTGTCTCT	260
DB	100	AGAGAAATGCAAGTACTCTTTGGAGGTTAAGAGTGTCTCTATGTCACTCTCTCTTT	159
QY	261	TGCTTGCCAAAGCACTTTTACAAATTTTGGGCGAGTTTACTATGAAATGGTGGGAAGAG	320
DB	160	AGCTGAGAGACAATAGCCACTCTCTTAGTAGACTTTACTACGAAAGGCGGTAAAG	219
QY	321	TACATGGATGGGAACACTTGTCCAACTAATCGGCTTCCCTGTCTGTCTTCTTCGCTT	380
DB	220	CACATGGCTCGAAACCTCTGGTTGAGCTTGTAGGGTTTCTTTAAACCTTCTCTGTATTA	279
QY	381	CTTTTCCCAACCAAAATCCCAACCAACAGACAGATTTTCAGAAAGTTCTCTTCCTT	440
DB	280	TTACTTAAAGCTTGAGCCGTCAAGACTAAACCAATTCACAAATAACTACTTCTTCCTT	339
QY	441	CACCAATCTTTGGATCAGTTTACATCGTTTACTGGACTATTAGTGTCTGTAACTCTTATAT	500
DB	340	CTTGACACTATCTTTAGTGTATATGGACTTGGCTTGGCTTGTGTGGACATTTGATTTT	399
QY	501	GTCTCTGTGTGTTTACTATACCTACAGTTTCTACTTTTCCCTCACTTGGCCTCACA	560
DB	400	GTACTCATTTGGGCTACTTTAGCTTCCCTCTCAACTTTCTCTTTGATCTCTGCGTCGA	459
QY	561	ATTGGCTTCACTGCTTCTTCTCATATTTTCTAACTCGGAGAGTTTCACACCTTTCAT	620
DB	460	ATTGGCTTTTAAAGCGGCTTCTCTTACTTCTTAACTCACAATAATCACACCACTTAT	519
QY	621	TGTGAATTTCTGTGTTCTCTTACTATTCTCTCTGCTTCTGCTTCTGCTGCTCAACTGATC	680
DB	520	ACTCAATTCACCTTGTTCTCTTAAACCATCTCTTACACTTCTTGTATTCACATGAACC	579
QY	681	GGAA-----AACACAGCAAAAGTGTCTAGAGTAAATATGTGATAGGATAAT	728
DB	580	AGAAATCTCCCTCTTCTTACTTCAAGTCCGAGCAAGTCCAAGTATGTGATTGGATACAT	639
QY	729	ATGTACCATTTGGTCTTCTGCTGGGATTTGGATTTGCTGCTATCCCTGGTACAACTGATCT	788
DB	640	CTGCGCGGTCGGTAGCTAGCTGGTGTATTTCTCTGCTGCTTCTTTTAAACAGATTACGCGTT	699
QY	789	CAGGAAGGTTTTTAAAGAACAACAACTTCTCAACGGTCACTGACTTGGTGGCTTTACCAATC	848
DB	700	CGAAAGATCTTAAGAAATACACATCTAAGGCTATTTTAGACATGGCCACATATCCGTC	759
QY	849	TCTAGTTGAAGCTGTGTGGTTCTCATAGGACTTTTTCGCAAGCGGGAGTGGAAACTTTT	908

DB	760	TATGTTAGTACTTCTGTGTAGTTGTGTTAGGACTTTTGGAGTGGTGGTGGAAAAAGCT	819
QY	909	AACRAAGTGAGATGGAATACTACAACTGGGAAAGTGCATACGTTATGACTTTTGGCCTC	968
DB	820	GAGTACAGAAATGGAAGAGTTTCAACTAGGAAAGCTCATACATTTTGTATAAATCGG	879
QY	969	GATAGCTATTTCTCGGCAAGTCTACACCAATGGCGTCTGGGACTGATCTTTGAGTCATC	1028
DB	880	TTCAACGATATCATGGCAAGCTTTGTTGATTGGAAGTGTGTTGATTATCGAAGTTTC	939
QY	1029	TTCTGTGTTCTCCAAATTCATAAATGCTGTGGGATTTGCTATAGTTCAGTTGTAGCAGT	1088
DB	940	ATCGCTTTTCTCCAAATGTCATAAGCACTCTTTGTTTACAGTTGTGCTCTTTCTGTGT	999
QY	1089	GATTCTTTTCCATGATAAAATGAACGCGTCAAGATCTTCTCCATCATTTTATGCTATCTG	1148
DB	1000	TGTCTTCTCGTATGAGTGAAGTGAATCAAGTTGATTGCAATGTTTGGCCATCTG	1059
QY	1149	GGGATTCATTTCTATTTGTTCTATCAGCACTACCTCGACGAAAGAAAG	1194
DB	1060	GGGATTTGTTCTTATGTTATCAGCAATATGTCATGATAGAAAG	1105
RESULT 12			
AAC47974			
ID	AAC47974 standard; DNA; 2660 BP.		
XX	AAC47974;		
XX	18-OCT-2000 (first entry)		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 55799.		
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
XX	promoter; termination sequence; ss.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
PD	06-SEP-2000.		
PF	25-FEB-2000; 2000EP-00301439.		
PR	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	30-APR-1999; 99US-0132407P.		
PR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	07-MAY-1999; 99US-0132487P.		
PR	11-MAY-1999; 99US-0132863P.		
PR	14-MAY-1999; 99US-0134218P.		
PR	14-MAY-1999; 99US-0134219P.		
PR	14-MAY-1999; 99US-0134221P.		
PR	14-MAY-1999; 99US-0134370P.		

PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139111P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140335P.
PR	24-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140931P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	06-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	26-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.

Db	743	CGGGCTCGTGGCGACGGCCGCTCGGTGTCGGGCTGTTTCGGAGCGGCGAGTGGAGGT	802
Qy	905	CTTTAACAGTGAAGTGAATAAAGTACAAACTGGGGAAAGTGCCATACGTTATGACTTTGG	964
Db	803	CGCTGAGGGGGAGATGAACGCGTTACAGTTCGGGCGAGTTCTCTACCTGATCAGCTGC	862
Qy	965	CCTCGATAGTATTTCTCTGGGAAGTCTACACCAATTGGCGTCGTGGGACTGATCTTTGAGT	1024
Db	863	TGTGGCGCGCGTGTCTGTGGCAGGTGGCCAACATCGGGGTGCTCGGCCCTCATCTTCGAGG	922
Qy	1025	CATCTTCTGTGTTCTCCAAATTCATAACTGCTGTGGGATGGCCTATAGTCCAGTTGTAG	1084
Db	923	TGTCGGCGCTCTTCTCCAAAGTATCAGCAGGTGTCACATGCGCGTCATCCCGTCTCTCG	982
Qy	1085	CAGTGATTTGTTTCCATGATAAAATGAACGCG	1116
Db	983	CGGTGGTGGTGTTCACAGCAGGATGAACGGG	1014

Search completed: November 1, 2004, 07:19:55
Job time : 684.972 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:44:43 ; Search time 5730.74 Seconds
(without alignments)
10669.751 Million cell updates/sec

Title: US-09-913-767-4

Perfect score: 1293
Sequence: 1 ggaagtcctcttagagtgtg.....tatgatcaaacatatattcc 1293

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_ste.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1293	100.0	1293	6	AX033547	AX033547 Sequence
2	1293	100.0	58427	8	ATF15J5	AL110123 Arabidops
3	1293	100.0	82697	8	ATT9A21	AL021713 Arabidops
4	1293	100.0	194143	8	ATCHRIV48	AL161548 Arabidops
5	1152	89.1	3387	6	AX507517	AX507517 Sequence
6	1152	89.1	3387	6	AX652015	AX652015 Sequence
7	816.6	63.2	1194	6	AX033548	AX033548 Sequence
8	775	59.9	1165	8	AY096558	AY096558 Arabidops
9	775	59.9	1428	8	AY074546	AY074546 Arabidops
10	454.4	35.1	1106	8	AF370622	AF370622 Arabidops
11	451.8	34.9	733	8	AF083706	AF083706 Arabidops
12	435.2	33.7	1071	6	AX033550	AX033550 Sequence
13	429.6	33.2	721	11	BX664674	BX664674 Arabidops
14	410	31.7	1047	6	AX412868	AX412868 Sequence
15	410	31.7	1081	6	AX033549	AX033549 Sequence
16	390.8	30.2	73179	8	AP006383	AP006383 Lotus cor
17	295	22.8	1617	8	AK072660	AK072660 Oryza sat
18	295	22.8	175645	8	AP004096	AP004096 Oryza sat
19	294.4	22.8	1145	6	AX033546	AX033546 Sequence

20	293.4	22.7	1476	8	AK071254	AK071254 Oryza sat
21	289	22.4	1251	8	BT003358	BT003358 Arabidops
22	289	22.4	1372	8	AY062782	AY062782 Arabidops
23	289	22.4	115421	8	T12C22	AC020576 Sequence
24	282.4	21.8	110036	8	AC142095	AC142095 Medicago
25	269	20.8	133299	8	OSJN00098	AL606658 Oryza sat
26	269	20.8	145739	8	OSJN00157	AL622957 Oryza sat
27	265.8	20.6	1486	8	AK119483	AK119483 Oryza sat
28	246.6	19.1	83698	8	AX540653	AX540653 Sequence
29	246.6	19.1	83698	8	AB010072	AB010072 Arabidops
30	245.2	19.0	1279	8	BT014359	BT014359 Lycopersi
c 31	234	18.1	104386	8	AT132A17	AL161813 Arabidops
c 32	234	18.1	199362	8	ATCHRIV24	AL161512 Arabidops
33	230.8	17.8	1117	8	AY096468	AY096468 Arabidops
34	230.8	17.8	1333	8	AY074535	AY074535 Arabidops
35	230.8	17.8	1403	8	AY087844	AY087844 Arabidops
36	227.2	17.6	2175	6	AX654201	AX654201 Sequence
37	214.4	16.6	1152	6	AX653666	AX653666 Sequence
38	179	13.8	1361	8	AK071680	AK071680 Oryza sat
39	157	12.1	1049	6	AX033545	AX033545 Sequence
40	140.6	10.9	1155	6	AX653139	AX653139 Sequence
41	140.6	10.9	1599	8	AK060241	AK060241 Oryza sat
42	128.4	9.9	1102	8	BT005504	BT005504 Arabidops
43	128.4	9.9	1250	8	AK117664	AK117664 Arabidops
44	126.8	9.8	1225	6	AX033544	AX033544 Sequence
45	126.8	9.8	1227	8	AF078531	AF078531 Arabidops

ALIGNMENTS

RESULT 1
AX033547
LOCUS AX033547 1293 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 4 from Patent WO0049152.
ACCESSION AX033547
VERSION AX033547.1 GI:10280291
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsais.
1
REFERENCE
AUTHORS Andre B., Buerkle, L., Frommer, W.B. and Gillissen, B.
TITLE Nucleic acids that code for a nucleobase transporter
JOURNAL Patent: WO 0049152-A 4 24-AUG-2000;
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;
GILLISSEN BERND (DE)

FEATURES
source
1..1293
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 100.0%; Score 1293; DB 6; Length 1293;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTCCTCTTAGAGTGGTCATGGAAATAACTCAAGTAATCTATGTCAATGGTAAGCAA 60
Db 1 GGAAGTCCTCTTAGAGTGGTCATGGAAATAACTCAAGTAATCTATGTCAATGGTAAGCAA 60

Qy 61 GATGCATCTCGTAGTGTAGATTACTTGTTCCTTAACCTGTTGTTTGTATGTTT 120
Db 61 GATGCATCTCGTAGTGTAGATTACTTGTTCCTTAACCTGTTGTTTGTATGTTT 120

Qy 121 TCAGATCATACACATAGAGCAACCTAACAGTCTAGGAGGAATAAATACCAATCCATG 180
Db 121 TCAGATCATACACATAGAGCAACCTAACAGTCTAGGAGGAATAAATACCAATCCATG 180

Qy 181 GAAATCGAATCTTCGTCGCTCACTCAATCGAAGAACTATAAGAAATGCTTCGTATTTC 240

Db	181	GAATCGAATCTTCTGTCGTAACCTCAATCGAAGAACTATAAGAAATGCGTTCGATTTC	240
Qy	241	ATTTACGTTCTTTGTCCTTCTGTCGTAACCTCAATCGAAGAAATGCGTTCGATTTC	300
Db	241	ATTTACGTTCTTTGTCCTTCTGTCGTAACCTCAATCGAAGAAATGCGTTCGATTTC	300
Qy	301	TATGAAATGTTGGGAAGAGTACATGGATGGGAACACTTGTCCAACTAATCGGCTTCCT	360
Db	301	TATGAAATGTTGGGAAGAGTACATGGATGGGAACACTTGTCCAACTAATCGGCTTCCT	360
Qy	361	GTTCCTGTTCTTCTGTCCTTCTTCCAAACCAAAATCCAAACCAAGAGAGAT	420
Db	361	GTTCCTGTTCTTCTGTCCTTCTTCCAAACCAAAATCCAAACCAAGAGAGAT	420
Qy	421	TTGAAAGTTCTCTTCTTCCACCAATCTTGGATCAGTTTACATCGTTACTGGACTATTA	480
Db	421	TTGAAAGTTCTCTTCTTCCACCAATCTTGGATCAGTTTACATCGTTACTGGACTATTA	480
Qy	481	GTGCTGTAACCTTATATGTCCTTCTGTTTACTATACCTACCGTTTCTACTTTC	540
Db	481	GTGCTGTAACCTTATATGTCCTTCTGTTTACTATACCTACCGTTTCTACTTTC	540
Qy	541	TCCCTCATCTGGCTCAAAATGGCCTTCACTGCTTCTTCTCATATTTCTAAACTCG	600
Db	541	TCCCTCATCTGGCTCAAAATGGCCTTCACTGCTTCTTCTCATATTTCTAAACTCG	600
Qy	601	CAGAAGTTACACCTTTCAATTTGTAATTTCTGTTTCTTCTTACTATTTCTGCGCTC	660
Db	601	CAGAAGTTACACCTTTCAATTTGTAATTTCTGTTTCTTCTTACTATTTCTGCGCTC	660
Qy	661	CTGTTGTTCAACCTGATTCGAAACACACAGAAAGTGTCTAGAGTAAATATGTGATA	720
Db	661	CTGTTGTTCAACCTGATTCGAAACACACAGAAAGTGTCTAGAGTAAATATGTGATA	720
Qy	721	GGGATAATGTACCATTTGTTCTTCTGCGATTTGATTTGCTATCCCTGCTACAA	780
Db	721	GGGATAATGTACCATTTGTTCTTCTGCGATTTGATTTGCTATCCCTGCTACAA	780
Qy	781	CTGATCTCTCAGGAAGTTTAAAGAACAAACATTTCTCAACGGTCACTGACTTGTGCT	840
Db	781	CTGATCTCTCAGGAAGTTTAAAGAACAAACATTTCTCAACGGTCACTGACTTGTGCT	840
Qy	841	TACCAATCTTAGTTGCAAGTGTGTTTCTCATAGGACTTTTTCGCAAGCGGAGTGG	900
Db	841	TACCAATCTTAGTTGCAAGTGTGTTTCTCATAGGACTTTTTCGCAAGCGGAGTGG	900
Qy	901	AAACCTTTAAACAAGTGAGATGAAACCTACAAACCTGGGAAAGTGCCATACGTTATGACT	960
Db	901	AAACCTTTAAACAAGTGAGATGAAACCTACAAACCTGGGAAAGTGCCATACGTTATGACT	960
Qy	961	TTGGCTCGATAGCTATTTCCTGGCAAGTCTACACCAATTTGCGTGGGACTGATCTTT	1020
Db	961	TTGGCTCGATAGCTATTTCCTGGCAAGTCTACACCAATTTGCGTGGGACTGATCTTT	1020
Qy	1021	GAGTCATCTTCTGTTCTTCAATTCATTAAGTGTGTTGGAATGCTATAGTTCCAGTT	1080
Db	1021	GAGTCATCTTCTGTTCTTCAATTCATTAAGTGTGTTGGAATGCTATAGTTCCAGTT	1080
Qy	1081	GTAGCAGTATTGTTTTCATGATAAATGAACCGGTCAAGAGATCTTCTCCATCATTTTA	1140
Db	1081	GTAGCAGTATTGTTTTCATGATAAATGAACCGGTCAAGAGATCTTCTCCATCATTTTA	1140
Qy	1141	GCTATCTGGGATTCATTTCTATGATGATAAATGAACCGGTCAAGAGATCTTCTCCATCATTTTA	1200
Db	1141	GCTATCTGGGATTCATTTCTATGATGATAAATGAACCGGTCAAGAGATCTTCTCCATCATTTTA	1200
Qy	1201	ACTAGCCACAAGTCTCTGATGAGATCTTCTATCTACTGCTGAGGAAGGTCACACA	1260
Db	1201	ACTAGCCACAAGTCTCTGATGAGATCTTCTATCTACTGCTGAGGAAGGTCACACA	1260
Qy	1261	AACATACATAGTATGATCAAAACATATTTCC	1293

Db	1261	AAATACATAGTATGATCAAAACATATTTCC	1293
RESULT 2			
ATF15J5			
LOCUS	ATF15J5	58427 bp	DNA linear
DEFINITION	Arabidopsis thaliana DNA chromosome 4, BAC clone F15J5 (ESSA project).		PLN 30-AUG-1999
ACCESSION	AL110123		
VERSION	AL110123.1	GI:5816989	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE			
AUTHORS	Bevan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H.W., Mayer, K.P.X., Lemcke, K. and Schueller, C.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 58427)		
AUTHORS	EU Arabidopsis sequencing project.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-MAR-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schuellemips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@brc.ac.uk		
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ .		
FEATURES			
source	Location/Qualifiers		
	1. .58427		
	/organism="Arabidopsis thaliana"		
	/mol_type="genomic DNA"		
	/variety="Columbia"		
	/db_xref="taxon:3702"		
	/chromosome="4"		
misc_feature	1. .2548		
	/note="overlap to BAC T6K21; please refer to EMBL:AL021889 for analysis and annotation"		
gene	3596. .5023		
	/gene="F15J5.10"		
gene	complement(join(3596. .3646,3736. .3801,3895. .4020,4116. .4281,4725. .5023))		
	/gene="F15J5.10"		
CDS	complement(join(3596. .3646,3736. .3801,3895. .4020,4116. .4281,4725. .5023))		
	/gene="F15J5.10"		
	/note="Contains Eukaryotic initiation factor 4E signature [DPYCFKHIEPKWEDPICANGKW] contains EST gb:T21435, AA586111"		
	/codon_start=1		
	/product="translation initiation factor eIF4E"		
	/protein_id="CA853645.1"		
	/db_xref="GI:5816990"		
	/db_xref="GOA:O23252"		
	/db_xref="Swiss-Prot:O23252"		
	/translation="MAVEDTPKSVTTEAKPNSIENPIDRYHEEGDDAEERGEIAGGEG DGNVDSKSGVPESHLESHWTFWFDNPAVKSKOTSGSSLRPVFTFSIVPEFWSLY NNKHKPSKLAHADFYCFKHIIEPKWEDPICANGKWTFPEKSKDSKMLYTLALI GEQFDHGEICGAVNVIRKQERISITWKNASNEAAQVSIQKQKWFELDYNNSIGFII HEDAKLDRNAKNAYTA"		
exon	complement(3596. .3646)		
	/gene="F15J5.10"		
	/number=1		
intron	complement(3647. .3735)		
	/gene="F15J5.10"		
	/number=1		
exon	complement(3736. .3801)		
	/gene="F15J5.10"		
	/number=2		

intron	complement (3802. .3894) /gene="F15J5.10" /number=2 complement (3895. .4020) /gene="F15J5.10" /number=3 complement (4021. .4115) /gene="F15J5.10" /number=3 complement (4116. .4281) /gene="F15J5.10" /number=4 complement (4282. .4724) /gene="F15J5.10" /number=4 complement (4725. .5023) /gene="F15J5.10" /number=5 6767. .6789 9077. .13984 /gene="F15J5.20" join(9077. .9301,9805. .9859,9938. .10113,10198. .10419, 10487. .10725,10806. .11331,11419. .11692,11793. .12736, 12827. .13093,13169. .13432,13508. .13984) /gene="F15J5.20" /note="similarity to multidrug resistance protein 1, Aspergillus fumigatus, U62933 Contains ABC transporters family signature. Abc Transporter [MSGGQKQRIAIARAI][LSGGQKQRIAIARAI], ATP/GTP-binding site motif A (P-loop), Atp_Gtp_A [GOSGSGKS][GESGSGKS]" /codon_start=1 /product="multidrug resistance protein/p-glycoprotein-like" /protein_id="CA853646.1" /db_xref="GI:5816991" /db_xref="GOA:Q9SVW6" /db_xref="TrEMBL:Q9SVW6" /translation="MEEKSKKNDGNQKVSFFKFLPSFADKTDVVLMTVGTIAAAGNG LTQPTWTLIFGQLINAFGTTDPDMVREVVKVAVFIYLAIVSCVVAFLQVSCWMTVG ERQSATIRGLYLKLTILRODIGYEDTETNGEVIKMSGDTILLIQDAMEKVKFTQLL CTFLGGFAIFYKGPLLAGVLCSCPLIVIAAASLIMSKWAGRQVAYAGNVVE QTVGAIRTVATYGEKQATEKYESKLEIAYKTVVQOGLISGFLGTMFLAVIFCSYGLA VNYGAKLIMEKNGYQGVINVI FAVLTGMSLQSTPSLNAFAAGRAAAAFKMFETIKR SPKIDAYDMGSLVEDIRGIDELKDVIFYPARPVDQIFPAGFSLFVPGKTVLALVQOS SGSKTVLSLIRFYPDESQVLIQNDLKLQKWLRSKIGLVSQEPVLEATTIKEN IAYGEDATDQIRTAIELANAAKFDKLPQGLDTHVGEHGTQMSGGQKQRIAIARAI LNPKILLDEATSAUDAESRIVQDALVNLMSNTTVVAHRLTTIRTDVIAVHHQ GKIVEKTHDEMIQDEGAYSQLVRLQEGSKEATESERPETSLDVERSGSLRLSSAM RRSVNRSSSRHSFSLASNPFPFGVNVNQTDEMEENNVRHKVSLKRLAHLNPE IPVLVLSIAAMVHGTVPFI FGLLLSSINMFVEPAKILKDSHFWALIYIALGNTF VMIPVNYFFAGGKLIKIRISMCFDQVHGOEISWFDNTANSRYNFIYINRRILY VLIILFICVLLPPLERECSDASTVRSVLGDALALIVQNTATVTTGLIIAFTANWI LALIIVLALSPFIVIOGYATKLTGFSADAKAMIEASQVANDAVSSIRTVASFCAEE KMYDYLQCDGKNGVRLGSLGAGFGFSFFLYCINCVCFVSGAGLIQIGKATFG EYFKVFALTIMAGIVSOTSAMAPSNKAKDAASTFOILDTPKIDSSDGTTLQNG PDGIEFRHFRYPMPRPVQIFRDLCLTIPSGKTVLALVGESEKSTVISMIERFYN VNSGKILLDOVEIRYQTFKLSLQQMLVSOEPILENTIRSNIAVYKGTGATEEIIA AKAANAHNFISSLFQGYDTSVGERGVQLSGGQKQRIAIARAILKDPKILLDEATSA LDAESERVVDALDRVXG"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
--------	---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 56690 TCAGATCAATAACAATGAAGCAAACTTAACAGTCAGGAGGAATGAATACCAACATG 56749
Qy 181 GAAATCGAATCTTCGTCGGTACCTCAATCGAAGAACTATAAGAAATGCTTCGTAATTC 240
Db 56750 GAAATCGAATCTTCGTCGGTACCTCAATCGAAGAACTATAAGAAATGCTTCGTAATTC 56809
Qy 241 ATTTACGTGTTCTTGTGCTTCTGCTTCCAAAGCACTTCTCAAAATTTGGGAGAGTTTAC 300
Db 56810 ATTTACGTGTTCTTGTGCTTCTGCTTCCAAAGCACTTCTCAAAATTTGGGAGAGTTTAC 56869
Qy 301 TATGAAATGTTGGGAGAGTACATGGAATGGAACACTTGTCCAACTAATCGGCTTCCCT 360
Db 56870 TATGAAATGTTGGGAGAGTACATGGAATGGAACACTTGTCCAACTAATCGGCTTCCCT 56929
Qy 361 GTTCTGTTCTTCTCCGCTTCTTTTCCAAACCAAAATCCAAACCAAGAGCAGAT 420
Db 56930 GTTCTGTTCTTCTCCGCTTCTTTTCCAAACCAAAATCCAAACCAAGAGCAGAT 56989
Qy 421 TTCAGAAAGTTCTTCTTCCCTTCCAACTTCTGGATCAGTTTACATCGTTACTGGACTATTA 480
Db 56990 TTCAGAAAGTTCTTCTTCCCTTCCAACTTCTGGATCAGTTTACATCGTTACTGGACTATTA 57049
Qy 481 GTGCTGCTAACTCTTATATGTCCTCTGTTGTTTACTATACTTACCAGTTTCTACTTTTC 540
Db 57050 GTGCTGCTAACTCTTATATGTCCTCTGTTGTTTACTATACTTACCAGTTTCTACTTTTC 57109
Qy 541 TCCCTCAFTCTGGCTCACAATGGCCCTTCTACTGCCCTTCTTCTCATATTTTCTAAACTCG 600
Db 57110 TCCCTCAFTCTGGCTCACAATGGCCCTTCTACTGCCCTTCTTCTCATATTTTCTAAACTCG 57169
Qy 601 CAGAAGTTACACCTTTCAATGTGAATCTCTGTTTCTCCTTACTATTTCTCTCGCCCTC 660
Db 57170 CAGAAGTTACACCTTTCAATGTGAATCTCTGTTTCTCCTTACTATTTCTCTCGCCCTC 57229
Qy 661 CTGCTGGTCAACACTGATTCGGAAACACACCAAAAGTCTAGAGTAAATATGTGATA 720
Db 57230 CTGCTGGTCAACACTGATTCGGAAACACACCAAAAGTCTAGAGTAAATATGTGATA 57289
Qy 721 GGGATAATATGATACCATTTGGTCTTCTGCTGGGATTTGGATTTGCTATCCCTCGGTACAA 780
Db 57290 GGGATAATATGATACCATTTGGTCTTCTGCTGGGATTTGGATTTGCTATCCCTCGGTACAA 57349
Qy 781 CTGATCTCAGAGAGTTTAAAGAGCAAAATCTCAACGGTCACTGACTTGGTCTGCT 840
Db 57350 CTGATCTCAGAGAGTTTAAAGAGCAAAATCTCAACGGTCACTGACTTGGTCTGCT 57409
Qy 841 TACCAATCTCTAGTTGCAAGCTGTGTGTTCTCTAGGACTTTTCGCAAGCGGGAGTGG 900
Db 57410 TACCAATCTCTAGTTGCAAGCTGTGTGTTCTCTAGGACTTTTCGCAAGCGGGAGTGG 57469
Qy 901 AAAAATTTAAACAGTGAGATGAAAACTACAACTGGGGAAAGTGCCATAGTTTATGACT 960
Db 57470 AAAAATTTAAACAGTGAGATGAAAACTACAACTGGGGAAAGTGCCATAGTTTATGACT 57529
Qy 961 TTGGCTCGATAGCTATTTCTTGGCAAGTCTACACCATTTGGCGTCTGGGACTGATCTTT 1020
Db 57530 TTGGCTCGATAGCTATTTCTTGGCAAGTCTACACCATTTGGCGTCTGGGACTGATCTTT 57589
Qy 1021 GAGTCATCTTCTGTGTTCTTCCAAATTCATAACTGCTGTGGGATTTGCCATAGTTTCCAGTT 1080
Db 57590 GAGTCATCTTCTGTGTTCTTCCAAATTCATAACTGCTGTGGGATTTGCCATAGTTTCCAGTT 57649
Qy 1081 GTACAGTGATTTGTTTCCATGATAAAATGAACCGTCAAAAGATCTTCTCCATCATTTTA 1140
Db 57650 GTACAGTGATTTGTTTCCATGATAAAATGAACCGTCAAAAGATCTTCTCCATCATTTTA 57709
Qy 1141 GCTATCTGGGATTTCAATTTGTTGTTATCAGCACTTACCTCGACGAAAGAGTTGAAG 1200
Db 57710 GCTATCTGGGATTTCAATTTGTTGTTATCAGCACTTACCTCGACGAAAGAGTTGAAG 57769
Qy 1201 ACTAGCACACAAGTCTCTGTAGGAGATCCTCATCTTACTACTCTGTAGGAGGTCACACA 1260

Db 57770 ACTAGCACACAAGTCTCTGTAGGAGATCCTCATCTACTACTGCTGAGGAGGTCACACA 57829
Qy 1261 AACATACATAGTGTATGATCAAAACATATATTTC 1293
Db 57830 AACATACATAGTGTATGATCAAAACATATATTTC 57862
RESULT 3
ATT9A21
LOCUS Arabidopsis thaliana DNA chromosome 4, BAC clone T9A21 (ESSA)
DEFINITION
ACCESSION AL021713
VERSION AL021713.1 GI:2832689
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1
AUTHORS Bevan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H.W., Mayer, K.P.X., Lemcke, K. and Schueller, C.
JOURNAL Unpublished
REFERENCE 2
AUTHORS (bases 1 to 82697)
JOURNAL EU Arabidopsis sequencing, project.
TITLE Direct Submission
SUBMITTED (22-SEP-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schueller@mps.biochem.mpg.de, mayer@mps.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
COMMENT Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.
FEATURES
source
1. 82697
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
402..1585
/gene="T9A21.10"
join(402..832,946..1090,1178..1585)
/gene="T9A21.10"
/notes="similarity to SPF1 protein, sweet protein, PIR2:S51529 and WRKY protein family, Petroselinum crispum, MNOS:S72443, MNOS:S72444, MNOS:S72445"
/codon_start=1
/product="DNA binding-like protein"
/protein_id="CAA16788.1"
/db_xref="GI:4038481"
/db_xref="GOA:O49720"
/db_xref="COA:Q8VWJ2"
/db_xref="Swiss-Prot:Q8VWJ2"
/translation="MSNETRDLYNQYPSFSLSHEMNLFTSPSSYGNLPSONGNP STYFSDCLQSSPAAYELQKTFGLSPSSSEVFNSSIDQEPNDRDVTNDVINGACNE TETRVSPNSSSSADHPGDSGKRRKRELVEEDQISKVGKTKKTEVKQREPRV SPTKSEVDHLEDGVRWKYGOKAVKNSPYPRIIANGNENRSYYRCTOKCNVKKRVE RSQDPTVITTYTGQHNHPIPTNLRGSSAAAMFSDLMTSPSFADHMERPTAAYNG GSVAAALDYGQSGYGSVNSPSSHQVTHQGGEYELLREIFFSIFFKQEP"
402..832
/gene="T9A21.10"
/number=1
833..945
/gene="T9A21.10"
/number=1
946..1090
/gene="T9A21.10"
/number=2
1091..1177
exon
intron
exon
intron

Db 8651 GATGATCTCGTAGTGATGATCTTTTCGTAACCTGTTGTTTGTGATGTTT 8710
Qy 121 TCAGATCATACAACATAGAGCAAACTTAACAGGTCAAGAGGAAATGAATACCAACATG 180
Db 8711 TCAGATCATACAACATAGAGCAAACTTAACAGGTCAAGAGGAAATGAATACCAACATG 8770
Qy 181 GAATCGAATCTTTCGTCGTAACCTCAATCGAGACTATAGAAATGCTTCGTAATTC 240
Db 8771 GAATCGAATCTTTCGTCGTAACCTCAATCGAGAACTATAAGAAATGCTTCGTAATTC 8830
Qy 241 ATTTACGTTCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 300
Db 8831 ATTTACGTTCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 8890
Qy 301 TATGAAATGTTGGGAGAGTACATGATGGAACACTTGTCCAACTAATCGGCTTCCT 360
Db 8891 TATGAAATGTTGGGAGAGTACATGATGGAACACTTGTCCAACTAATCGGCTTCCT 8950
Qy 361 GTTCGTTCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 420
Db 8951 GTTCGTTCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 9010
Qy 421 TTCAGAAATGTTCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 480
Db 9011 TTCAGAAATGTTCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 9070
Qy 481 GTGCTGTAATCTTATATGTCCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 540
Db 9071 GTGCTGTAATCTTATATGTCCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 9130
Qy 541 TCCCTCATCTTGGGCTCAATGTCCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 600
Db 9131 TCCCTCATCTTGGGCTCAATGTCCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 600
Qy 601 CAGAGTTTCAACACTTTCGTCGTAACCTCAATGTCCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 660
Db 9191 CAGAGTTTCAACACTTTCGTCGTAACCTCAATGTCCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 9250
Qy 661 CTCGGTCAACACTTTCGTCGTAACCTCAATGTCCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 720
Db 9251 CTCGGTCAACACTTTCGTCGTAACCTCAATGTCCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 9310
Qy 721 GGGATAATGATACATGTCCTTTCGTCGTAACCTCAATGTCCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 780
Db 9311 GGGATAATGATACATGTCCTTTCGTCGTAACCTCAATGTCCTTTCGTCGTAACCTCAATCGAGCACTTTCTACAATTTTGGGAGAGTTTAC 9370
Qy 781 CTGATCCTCAGAGGTTTAAAGAGCAAACTTCTCAACGGTCACCTGACTTGGTCGCT 840
Db 9371 CTGATCCTCAGAGGTTTAAAGAGCAAACTTCTCAACGGTCACCTGACTTGGTCGCT 9430
Qy 841 TACCAATCTTAGTTGCAAGTGTGTGTTCTCATAGGACTTTTCGCAAGCGGAGTGG 900
Db 9431 TACCAATCTTAGTTGCAAGTGTGTGTTCTCATAGGACTTTTCGCAAGCGGAGTGG 9490
Qy 901 AAACCTTAAACAGTACAGTGGAAACCTCAAACTGGGAAAGTGCCTATGACT 960
Db 9491 AAACCTTAAACAGTACAGTGGAAACCTCAAACTGGGAAAGTGCCTATGACT 9550
Qy 961 TTGGCTCGATAGCTATTTCTCGCAAGTGTACACATTTGGCGTGGGACTGATCTTT 1020
Db 9551 TTGGCTCGATAGCTATTTCTCGCAAGTGTACACATTTGGCGTGGGACTGATCTTT 9610
Qy 1021 GAGTCATCTTCTGTGTTCTCAATTTCCAAATACCTGCTGGGAGTGCCTATGACTTCCAGTT 1080
Db 9611 GAGTCATCTTCTGTGTTCTCAATTTCCAAATACCTGCTGGGAGTGCCTATGACTTCCAGTT 9670
Qy 1081 CTAGCAGTATGTTTTCATGATATAATGACCGGTCAAGAGTCTTCTCCATCATTTTA 1140
Db 9671 CTAGCAGTATGTTTTCATGATATAATGACCGGTCAAGAGTCTTCTCCATCATTTTA 9730
Qy 1141 GCTATCTGGGATTCATTTTCTATCTAGCACTACCTCGACGAAAGAGTTTGAAG 1200

Db 9731 GCTATCTGGGATTCATTTTCTATCTAGCACTACCTCGACGAAAGAGTTTGAAG 9790
Qy 1201 ACTAGCCACACAAGTCTCTAGGAGATCCTCATCTACTACTCTGCTAGGAGGTCACACA 1260
Db 9791 ACTAGCCACACAAGTCTCTAGGAGATCCTCATCTACTACTCTGCTAGGAGGTCACACA 9850
Qy 1261 AACATACATAGTGTATGATCAAAACATATTTCC 1293
Db 9851 AACATACATAGTGTATGATCAAAACATATTTCC 9883
RESULT 4
ATCHRIV48
LOCUS ATCHRIV48 194143 bp DNA linear PLN 16-MAR-2000
DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 48.
ACCESSION AL161548
VERSION AL161548.2 GI:7268604
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 111084)
MURPHY, G., RIDLEY, P., HUDSON, S., MEWES, H.W., LEMCKE, K. and MAYER, K.F.X.
Unpublished
REFERENCE 2 (bases 107966 to 194143)
HILBERT, H., BRAUN, M., HOLZER, E., BRANDT, A., DUESTERHOEF, A., MEWES, H.W., LEMCKE, K. and MAYER, K.F.X.
Unpublished
JOURNAL 3 (bases 1 to 194143)
EU Arabidopsis sequencing, project.
REFERENCE Direct Submission
AUTHORS Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
COMMENT Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/> this fragment has an overlap with ATCHRIV47 at the 5' end and an overlap with ATCHRIV49 at the 3' end.
FEATURES
Location/Qualifiers
1. 194143
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
6377..9293
/gene="AT4g18120"
join(6377..6796,7122..7238,7347..7501,7578..8456,8543..8627,8716..8784,8871..9293)
/gene="AT4g18120"
/notes="similarity to Me12-like protein, Arabidopsis thaliana, gb:D86122 contains EST gb:W43784, AI998063.1"
/codon_start=1
/product="putative protein"
/protein_id="CAB78814.1"
/db_xref="GI:7268605"
/db_xref="GOA:Q9SVV9"
/db_xref="TrEMBL:Q9SVW9"
/translation="MASGLDFHSGIGNMLDGDHPHPIGNMLPDEBELFSGMLDDLNL LSSIPATLDLEDYDLFGSGGLELTDPYDSLNGKFSRMGFADSNVDNVPQNI FQNVGVSGIAGEHYGHPNTLFRVNRINSVESELQALPEDNPSKVDNQGLVFNPLAP SVNRDLNIFGVYGEIKETRETPNKRHKHFEFFDVRSDAALKALNRETAKKRI KLHRSRPGARRNMQLQNPEDDDSYSLNHFVEPLASSPIGKWNNSPIDPLQSP SKSPIFGNLSPTKNIRYPEFSMTKASVYNDOEGRRFSLDLHFLSSSYNNASHKASTP QPQSGSVSSFGSLNSHPHVETLSGSEFLWGSPPSSAWPNVPFSSNRENHRFPYSA

QNGSLHQLHHIGSAPSGFFPRSPETSSMGSAVPRGASGNMNAQNRNLRRTSSNPNFMLS
APRSQULTGNGSYLMPAAATMWSIDDDPBGSGNQDDSGNQADIKIQQLDLSKIMR
GEDPRTLMIKNIPKNTYRNMLLAAIDERNSTGYDFLYLPIDFKNKNVGYAFINNV5
PTTIALYFAFNGKWDKFNSEKVASLAVRQGAALIAHFQNSLNMNDRRCOPIV
FDGSESKYPWNLLQOHTDLKMIYILFCVYLMNWLHTCLFETRSGGEFAARASNSTG
FLIFLSARTHIKADSQNVHVIASGLKRRKQFLG"
6377..6796

exon
/gene="AT4g18120"
/number=1

intron
6797..7121

exon
/gene="AT4g18120"
/number=1

intron
7122..7238

exon
/gene="AT4g18120"
/number=2

intron
7239..7346

exon
/gene="AT4g18120"
/number=2

intron
7347..7501

exon
/gene="AT4g18120"
/number=3

intron
7502..7577

exon
/gene="AT4g18120"
/number=3

intron
7578..8456

exon
/gene="AT4g18120"
/number=4

intron
8457..8542

exon
/gene="AT4g18120"
/number=4

intron
8543..8627

exon
/gene="AT4g18120"
/number=5

intron
8628..8715

exon
/gene="AT4g18120"
/number=5

intron
8716..8784

exon
/gene="AT4g18120"
/number=6

intron
8785..8870

exon
/gene="AT4g18120"
/number=6

intron
8871..9293

exon
/gene="AT4g18130"
/number=7

gene
9594..13230

gene
complement(join(9594..9837,9944..10237,10323..11130,
11238..13230))

CDS
/gene="AT4g18130"
complement(join(9594..9837,9944..10237,10323..11130,
11238..13230))

exon
/gene="AT4g18130"
/note="Contains ATP/GTP-binding site motif A (P-loop)
AA634-641;Phytochrome chromophore attachment site
AA317-326

intron
contains EST gb:T21020, AA395462"
/codon_start=1

gene
/product="phytochrome E"
/protein_id="CAB78815.1"

gene
/db_xref="GI:7268606"
/db_xref="GOA:P42498"

gene
/db_xref="Swiss-Prot:P42498"

gene
/translation="MGFSSSSAASNMKPOKSNATQYSDAALFADPAOSIVTCKS
FNYKSIVSPNHVDEHITAYLSNIQRGLVQPGCLIAVEEPSFRILGLSDNSDF
LGLSLPSTSGSEGDVKVGLIGIDARTLFTPTSPGASLSKAAASFTIISLNPVHVS
TTQKPFYALHRIADIGVMDLEPAKSGDPALTAGVQSKLAVRAISRLOSPLGDI
GALCDTVDRVQVQYQFHEDDHGEVSEIRSDLEPYLGLHYPATDIPQA
ARFLPKQNRVMI CDNATPVKVQSEELRPLCLVNSTLRAPHCHQTYMANMGSA
SLALAIVVKGDSSKLWGLVGHGCSPRVVPFLVACEFLMQAFGLQLOLQWELQASQ
LAEKAMRTQTLCDMLLRDTSVAIVTQSPGIMDLVKCDGGAALYTKGKCWLIVGVTPE
SQVKDLVNWLENHGDSDGLTFLTDSLVDAGYPGAI SLGDVCGVAAAFSSKDYLLWF
RSNTASAIKGGAKHPKDKODAGRMRIPRSFTAFLEVAKRSRLPWEISEIDAHSRL

LIMRESFTSSRPVLSGNGVARDANBELTSFVCEMVRVITATAPIFGVDSSGCGINGWK
KTAENTGLLASEAMGKSLADEIVQESRAALESILCKALQGEESKVMKLRKFGQNN
HPDYSDDCVLVNSCTSRDYTENIIGVCFVGODITSEKAITDRFIRLQCDYKTIYQSL
NPLIPPIASDENACCSWNAAMEKLTGMSKHEVIGKMLPGEVFGVCKVKQDSLTAK
FLISLYQIAGDNVPESSELIVFEFFNKGKYLEASLTANKSTNIEGKIVRCFFPFIQIINK
ESGLSCPBLKSAQSLNELTYVRQIKNPLNGIRFAHKLLESSEISASORQFLETSDA
CEKQITTIESTDLKSI BEGKLQLETFRELIENILDTTISQVMI LIRENSQRUEVA
EEIKTTLPLNGDRVKQLILADLRNIVNHAFFPNPSWVGTSISPGQELSRDNGRYIHQ
FRMIHPGKGLPSELSMDMFEETRDMGVTDPGLGLKLSRKILLEQMNGRVS YVREDERCFF
QVDLQVKMLGVESRGTRGSSSIK"

exon
complement(9594..9837)
/gene="AT4g18130"
/number=1

intron
complement(9838..9943)
/gene="AT4g18130"

exon
/number=1
complement(9944..10237)
/gene="AT4g18130"

intron
/number=2
complement(10238..10322)
/gene="AT4g18130"

exon
/number=2
complement(10323..11130)
/gene="AT4g18130"

intron
/number=3
complement(11131..11237)
/gene="AT4g18130"

exon
/number=3
complement(11238..13230)
/gene="AT4g18130"

gene
14593..16312
/gene="AT4g18140"

gene
complement(join(14593..14751,14921..15070,15338..15639,
15915..16055,16141..16312))

CDS
/gene="AT4g18140"
complement(join(14593..14751,14921..15070,15338..15639,
15915..16055,16141..16312))

exon
/gene="AT4g18140"
/note="similarity to HYA22 protein, Homo sapiens,
PIR2:JCS707

contains EST gb:R65235"
/codon_start=1

gene
/product="putative protein"
/protein_id="CAB78816.1"

gene
/db_xref="GI:7268607"
/db_xref="TrEMBL:Q9SVV8"

gene
/translation="MKSISKISLREKRVLGVCKPKHDKVNSCFDPKISERMELLVT
ATQCNNEEQNIGLDSRDVTNEPQDLVEDSSKNASNNMTIFSPVLDDLYSET
DREPSVLIDAAEQYMLLPYLRDRSASSDDVKSDENRINRSRKNLEAAENHTEAQT
EDFDPQIPLRNQPELADVFNYPVDMQOPRDSPKRKA VTFVLIWMLLDILOPDGKFS
QRFYRDSCLISDGIYTKDYLVLGLDLAKVALVDNCPQYVRLQINNGIFIKSWYDDPTD
DGLTLLPFLLETLDANDVRPIAKRFQNK"

exon
complement(14593..14751)
/gene="AT4g18140"

intron
/number=1
complement(14752..14920)
/gene="AT4g18140"

exon
/number=1
complement(14921..15070)
/gene="AT4g18140"

intron
/number=2
complement(15071..15337)
/gene="AT4g18140"

exon
/number=2
complement(15338..15639)
/gene="AT4g18140"

Query Match
Best Local Similarity 100.0%; Score 1293; DB 8; Length 194143;
Matches 1293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 36978 GGAAGTCTCTTAGAGTGGTCATGGAAATACTCAAGTAATCTATGTCAATGGTAAGCAA 37037
Qy 61 GATGATCTCGTAGTGTAGATTACTTTGATTTCTTTTCGCTAACCTGTGTTGTTTTCATGTTT 120
Db 37038 GATGATCTCGTAGTGTAGATTACTTTGATTTCTTTTCGCTAACCTGTGTTTTCATGTTT 37097
Qy 121 TCAGATCATACACATAGAGCAAACTTAACAGGTACAGGAGGAAAATGAATACCACCATG 180
Db 37098 TCAGATCATACACATAGAGCAAACTTAACAGGTACAGGAGGAAAATGAATACCACCATG 37157
Qy 181 GAAATCGAATCTTCGTCGTCCTCAATCGAAGAACTATAAGAAATGCTTCGTTATTTCC 240
Db 37158 GAAATCGAATCTTCGTCGTCCTCAATCGAAGAACTATAAGAAATGCTTCGTTATTTCC 37217
Qy 241 ATTTACGTTCTTTTGTCTCTTTCGTCGCAAGCACCTTTCTACAATTTTGGGCAGAGTTTAC 300
Db 37218 ATTTACGTTCTTTTGTCTCTTTCGTCGCAAGCACCTTTCTACAATTTTGGGCAGAGTTTAC 37277
Qy 301 TATGAAATGGTGGGAAGATACATGGATGGGAACACTTGTGCCAATTAATCGGCTTCCT 360
Db 37278 TATGAAATGGTGGGAAGATACATGGATGGGAACACTTGTGCCAATTAATCGGCTTCCT 37337
Qy 361 GTTCGTTCTCTTCGCTCTCTTTCCCAACCAAAATCCAAACCAAGAGAGAGAT 420
Db 37338 GTTCGTTCTCTTCGCTCTCTTTCCCAACCAAAATCCAAACCAAGAGAGAGAT 37397
Qy 421 TTCAGAAAGTTCTCTTCCTCCACCATCTTCGATCAGTTTACATCGTTACTGGAATA 480
Db 37398 TTCAGAAAGTTCTCTTCCTCCACCATCTTCGATCAGTTTACATCGTTACTGGAATA 37457
Qy 481 GTGTCGTAACTCTTATATGTCCTCTGTTGGTTTACTATTAATACAGTTTCTACTTTTC 540
Db 37458 GTGTCGTAACTCTTATATGTCCTCTGTTGGTTTACTATTAATACAGTTTCTACTTTTC 37517
Qy 541 TCCCTCATCTTGGCTCACAATTTGGCTTCTACTGCTTCTCTCATATTTTCTAAACTCG 600
Db 37518 TCCCTCATCTTGGCTCACAATTTGGCTTCTACTGCTTCTCTCATATTTTCTAAACTCG 37577
Qy 601 CAGAAAGTTACACCTTTTCATTTGTAATTTCTGTTTTCCTTTACTATTTCTCGCCCTC 660
Db 37578 CAGAAAGTTACACCTTTTCATTTGTAATTTCTGTTTTCCTTTACTATTTCTCGCCCTC 37637
Qy 661 CTCGTGTTCAACACTGATTCGGGAACACACGAAAGTGTCTAGAGTAATATGTGATA 720
Db 37638 CTCGTGTTCAACACTGATTCGGGAACACACGAAAGTGTCTAGAGTAATATGTGATA 37697
Qy 721 GGGATAATATGTACCATTTGGTCTTCTGCTGGGATTGGATTGCTGCTATCCCTGGTACAA 780
Db 37698 GGGATAATATGTACCATTTGGTCTTCTGCTGGGATTGGATTGCTGCTATCCCTGGTACAA 37757
Qy 781 CTGATCTCAGAAAGTTTAAAGAACAAACATTTCTCAACGGTCACGTGCTTGGTCGCT 840
Db 37758 CTGATCTCAGAAAGTTTAAAGAACAAACATTTCTCAACGGTCACGTGCTTGGTCGCT 37817
Qy 841 TACCAATCTTAGTTGCAAGTGTGTGTTTCTCATAGACATTTTCGCAAGGGGGAGTGG 900
Db 37818 TACCAATCTTAGTTGCAAGTGTGTGTTTCTCATAGACATTTTCGCAAGGGGGAGTGG 37877
Qy 901 AAAACTTTAAACAAGTGAGATGGAAAATPACAAAATCTGGGGAAGTGGCCATAGTTTATGACT 960
Db 37878 AAAACTTTAAACAAGTGAGATGGAAAATPACAAAATCTGGGGAAGTGGCCATAGTTTATGACT 37937
Qy 961 TTGGCCTCGATAGCTATTTCTTCGCAAGTCTPACACCATTTGGCGTGGGACTGATCTTT 1020
Db 37938 TTGGCCTCGATAGCTATTTCTTCGCAAGTCTPACACCATTTGGCGTGGGACTGATCTTT 37997
Qy 1021 GAGTCACTCTTGTTGTTCTCCAAATTCATTAATCTGCTGGGATTCGCTATAGTTCCAGTT 1080
Db 37998 GAGTCACTCTTGTTGTTCTCCAAATTCATTAATCTGCTGGGATTCGCTATAGTTCCAGTT 38057
Qy 1081 GTAGCAGTGATTGTTTTCATGATATAAATGAACCGGTCAAGAGATCTTCTCCATCATTTTA 1140

Db 38058 GTACAGTGTATGTTTTCATGATATAAATGAACGGTCAAGATCTTCTCCATCATTTTA 38117
Qy 1141 GCTATCTGGGGATTCATTTCAATTTGTTCTATCAGACCTACCTCGACGAAAAGAGTTGAAG 1200
Db 38118 GCTATCTGGGGATTCATTTCAATTTGTTCTATCAGACCTACCTCGACGAAAAGAGTTGAAG 38177
Qy 1201 ACTAGCCACACAACCTCTGCTAGGAGATCCTCATCTACTCTCTGCTAGGAGGTCACACA 1260
Db 38178 ACTAGCCACACAACCTCTGCTAGGAGATCCTCATCTACTCTGCTAGGAGGTCACACA 38237
Qy 1261 AACATACATAGTGTATGATCAAAACATATTTCC 1293
Db 38238 AACATACATAGTGTATGATCAAAACATATTTCC 38270
RESULT 5
AX507517 3387 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 2212 from Patent WO0216655.
DEFINITION AX507517
ACCESSION AX507517
VERSION AX507517.1 GI:23388754
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1
Harper, J.F., Krepes, J., Wang, X. and Zhu, T.
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
Patent: WO 0216655-A 2212 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source 1. .3387
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 89.1%; Score 1152; DB 6; Length 3387;
Best Local Similarity 96.4%; Pred. No. 7.6e-286;
Matches 1207; Conservative 0; Mismatches 0; Indels 45; Gaps 1;
Qy 22 ATGGAATAAATCTCAAGTAATCTATGTCAATGGTAAGCAAGATGATCTCGTAGTGTAGAT 81
Db 1 ATGGAATAAATCTCAAGTAATCTATGTCAATGGTAAGCAAGATGATCTCGTAGTGTAGAT 60
Qy 82 TACTTGATTTCTTTTCGCTAACCTGTTGTTTTCGATGTTTTCAGATCATACACATAGAA 141
Db 61 TACTTGATTTCTTTTCGCTAACCTGTTGTTTTCGATGTTTTCAGATCATACACATAGAA 120
Qy 142 GCAAACTCAACAGGTACAGGAGGAAATGAATACCAACCATGGAATCGAATCTTCGTCGTA 201
Db 121 GCAAACTCAACAGGTACAGGAGGAAATGAATACCAACCATGGAATCGAATCTTCGTCGTA 180
Qy 202 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTTCCATTTACGTTGTTCTTTGTCCTT 261
Db 181 CCTCAATCGAAGAACTATAAGAAATGGCTTCGTAATTTCCATTTTACGT----- 227
Qy 262 GCTTGCCAAAGCATTTTCTACAAATTTGGGCAGAGTTTACTATGAATAATGGTGGGAAGGT 321
Db 228 -----AGTTTACTATGAATAATGGTGGGAAGGT 255
Qy 322 ACATGGATGGGAACACCTTGTCCAACTAATCGGCTTCCTCTGTTCTCTTCCTCCGCTTC 381
Db 256 ACATGGATGGGAACACCTTGTCCAACTAATCGGCTTCCTCTGTTCTCTTCCTCCGCTTC 315
Qy 382 TTTTCCCAACCCAAAAATCCAAACCAAGAGAGAGATTTTCAGAAAGTTCTCTTCCTTC 441
Db 316 TTTTCCCAACCCAAAAATCCAAACCAAGAGAGATTTTCAGAAAGTTCTCTTCCTTC 375

QY	442	ACAACTCTGGATCAGTTTACATCGTTACTGGACTAATAGTGTCTGCTAACTCTTATATG	501
Db	376	ACAACTCTGGATCAGTTTACATCGTTACTGGACTAATAGTGTCTGCTAACTCTTATATG	435
QY	502	TCCTCTGTTGGTTTACTATATCTACAGTTTCTACCTCTCTCCCTCATCTTGGCCCTCACAA	561
Db	436	TCCTCTGTTGGTTTACTATATCTACAGTTTCTACCTCTCTCCCTCATCTTGGCCCTCACAA	495
QY	562	TTGGCCCTTCACTGCTCTTCTCATATTTTCTAACTCGCAGAGTTTCAACCTTTTCATT	621
Db	496	TTGGCCCTTCACTGCTCTTCTCATATTTTCTAACTCGCAGAGTTTCAACCTTTTCATT	555
QY	622	GTGAATTTCTGTTTCTCTTACTATTTCTCTGCTCTCTCTGCTCTCTCTGCTCAACATGCTG	681
Db	556	GTGAATTTCTGTTTCTCTTACTATTTCTCTGCTCTCTCTGCTCTCTCTGCTCAACATGCTG	615
QY	682	GAACAACACAGCAAGTGTCTAGTAAATATGTATAGGATATATGTATACCAATGGT	741
Db	616	GAACAACACAGCAAGTGTCTAGTAAATATGTATAGGATATATGTATACCAATGGT	675
QY	742	GCTTCTGCTGGATTTGGATTTGCTTATCCCTGCTCAAACTGATCTCTCAGGAAGTTTAA	801
Db	676	GCTTCTGCTGGATTTGGATTTGCTTATCCCTGCTCAAACTGATCTCTCAGGAAGTTTAA	735
QY	802	AGAAGCAAACTTCTCAACGGTCACTGACTGCTGCTGCTTCAAACTCTCTAGTTGCAAGC	861
Db	736	AGAAGCAAACTTCTCAACGGTCACTGACTGCTGCTGCTTCAAACTCTCTAGTTGCAAGC	795
QY	862	TGTGTGTTCTCATAGGACTTTTCCGACGGGGAGTGGAACCTTAACTAGTGAGTG	921
Db	796	TGTGTGTTCTCATAGGACTTTTCCGACGGGGAGTGGAACCTTAACTAGTGAGTG	855
QY	922	GAACAACACAGCAAGTGTCTAGTAAATATGTATAGGATATATGTATACCAATGGT	981
Db	856	GAACAACACAGCAAGTGTCTAGTAAATATGTATAGGATATATGTATACCAATGGT	915
QY	982	TGCGAAGTCTACACCAATCGGCTGCTGGAGTGAATCTTTGAGTCACTCTCTGTTCTCC	1041
Db	916	TGCGAAGTCTACACCAATCGGCTGCTGGAGTGAATCTTTGAGTCACTCTCTGTTCTCC	975
QY	1042	AAATCCATCACTGCTGGATTTGCTATAGTTTCCAGTTGTAGCAGTGAATTTTCCAT	1101
Db	976	AAATCCATCACTGCTGGATTTGCTATAGTTTCCAGTTGTAGCAGTGAATTTTCCAT	1035
QY	1102	GATAAAATGAACGGCTCAAGATCTTCTCCATCATTTTATAGTATCTGGGATTTCAATTC	1161
Db	1036	GATAAAATGAACGGCTCAAGATCTTCTCCATCATTTTATAGTATCTGGGATTTCAATTC	1095
QY	1162	TTTGTCTATCAGCACTTACCTCGACGAAAGAGTTGAAGACTAGCCACAAAGTCTGTGA	1221
Db	1096	TTTGTCTATCAGCACTTACCTCGACGAAAGAGTTGAAGACTAGCCACAAAGTCTGTGA	1155
QY	1222	CGAGATCCCTACTACTCTGAGGAGGCTCAGCAAAATACATAGTG	1273
Db	1156	CGAGATCCCTACTACTCTGAGGAGGCTCAGCAAAATACATAGTG	1207

RESULT 6
AX652015
LOCUS AX652015 3387 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 908 from Patent WO03000898.
ACCESSION AX652015
VERSION AX652015.1 GI:29154833

KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katagiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE Plant genes involved in defense against pathogens

JOURNAL	Patent: WO 03000898-A 908 03-JAN-2003;									
FEATURES	Syngenta Participations AG (CH)									
source	Location/Qualifiers									
	1..3387									
	/organism="Arabidopsis thaliana"									
	/mol_type="unassigned DNA"									
	/db_xref="taxon:3702"									
ORIGIN										
	Query Match	89.1%;	Score 1152;	DB 6;	Length 3387;					
	Best Local Similarity	96.4%;	Pred. No. 7.6e-286;							
	Matches 1207;	Conservative 0;	Mismatches 0;	Indels 45;	Gaps 1;					
QY	22	ATGGAATAA	CTCAAGTAATCTATGTC	CAATG	GTAAAGCAAGATGC	ATCTCTG	TAGT	TAGAT	81	
Db	1	ATGGAATAA	CTCAAGTAATCTATGTC	CAATG	TGGTAAGCAAGATGC	ATCTCTG	TAGT	TAGAT	60	
QY	82	TACTTGATTTCTTT	CGCTAACTGTTGTTT	GTATG	TTTTGATGTTTT	CAGATCAT	AACAACATAGAA	141		
Db	61	TACTTGATTTCTTT	CGCTAACTGTTGTTT	GTATG	TTTTGATGTTTT	CAGATCAT	AACAACATAGAA	120		
QY	142	GCAAACTAACAGGT	CAGGAGGAAATGAAT	ACCA	CCATGGAAATCGAA	ATCTTCG	TCGCGTA	201		
Db	121	GCAAACTAACAGGT	CAGGAGGAAATGAAT	ACCA	CCATGGAAATCGAA	ATCTTCG	TCGCGTA	180		
QY	202	CCTCAATCGAAGAA	CTATAAGAAATGGCTTC	GTATTTCC	ATTTACGTTGTTCTTT	GTCCCTT	261			
Db	181	CCTCAATCGAAGAA	CTATAAGAAATGGCTTC	GTATTTCC	ATTTACGTTGTTCTTT	GTCCCTT	227			
QY	262	GCTTGCCAAAGCA	CTTTCTACAATTTT	GGGCAGAGTTTT	ACTATGAAAATGGTGGGAAG	AGT 321				
Db	228	-----	-----	-----	-----AGTTTACTATGAAAATGGTGGGAAG	AGT 255				
QY	322	ACATGGATGGGAA	CACTTGTCCAACTAAT	CGGCTTCCCT	TGTTCTGTTTCTCTTCCGCTTC	381				
Db	256	ACATGGATGGGAA	CACTTGTCCAACTAAT	CGGCTTCCCT	TGTTCTGTTTCTCTTCCGCTTC	315				
QY	382	TTTTCCCAAAACCA	AAATCCCAACCAAG	CAGAGCAGATTT	CAGAAAGTTCTCTTCCCTTC	441				
Db	316	TTTTCCCAAAACCA	AAATCCCAACCAAG	CAGAGCAGATTT	CAGAAAGTTCTCTTCCCTTC	375				
QY	442	ACCATTTCTTGGAT	CAGTTTACATCGTT	TACTGGA	CTATTAGTGTCTGCTAACTCTTATATG	501				
Db	376	ACCATTTCTTGGAT	CAGTTTACATCGTT	TACTGGA	CTATTAGTGTCTGCTAACTCTTATATG	435				
QY	502	TCCTCTGTTGGTT	TACTATGTTTACAG	TTTCTACTTTCT	CCCTCATCTTGGCCTCACAA	561				
Db	436	TCCTCTGTTGGTT	TACTATGTTTACAG	TTTCTACTTTCT	CCCTCATCTTGGCCTCACAA	495				
QY	562	TTGGCCTTCACTG	CTTCTTCTCATATTTT	CTAAACTCGCAG	AAAGTTTCAACCTTTTCATT	621				
Db	496	TTGGCCTTCACTG	CTTCTTCTCATATTTT	CTAAACTCGCAG	AAAGTTTCAACCTTTTCATT	555				
QY	622	GTGAATTTCTGTT	TCTCTTACTATTTT	CTCTG	CCCTCCTCTGCTGCTCAACACTGATTCG	681				
Db	556	GTGAATTTCTGTT	TCTCTTACTATTTT	CTCTG	CCCTCCTCTGCTGCTCAACACTGATTCG	615				
QY	682	GAAACACAGCAAA	AGTGTCTAGAGTAA	ATAATGTGATAGG	ATAATATGTACCATTTGGT	741				
Db	616	GAAACACAGCAAA	AGTGTCTAGAGTAA	ATAATGTGATAGG	ATAATATGTACCATTTGGT	675				
QY	742	GCTTCTGCTGGAT	TTGGATTTGCTGCT	TATCCCTGGT	TAACTGATCTCAGGAAGGTTTTA	801				
Db	676	GCTTCTGCTGGAT	TTGGATTTGCTGCT	TATCCCTGGT	TAACTGATCTCAGGAAGGTTTTA	735				
QY	802	AAGAAGCAAACT	TCTCAACGGTCACT	GACTGCTTGGT	CGCTTACCAATCTCTAGTTGCAAGC	861				
Db	736	AAGAAGCAAACT	TCTCAACGGTCACT	GACTTGGT	CGCTTACCAATCTCTAGTTGCAAGC	795				
QY	862	TCGTGTGTTCT	CATAGGACTTTT	TCGACGGGGAGT	GGAAACCTTAACTAGTGAGATG	921				
Db	796	TCGTGTGTTCT	CATAGGACTTTT	TCGACGGGGAGT	GGAAACCTTAACTAGTGAGATG	855				

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE AUTHORS

1 (bases 1 to 1165)
Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished

TITLE JOURNAL REFERENCE AUTHORS

2 (bases 1 to 1165)
Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission

TITLE JOURNAL COMMENT

Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN
Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members constructed and
sequenced the pUNI (ORF) clones using the RAPL cDNAs: Yamada, K.,
Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,
Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,
Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P.,
Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally
to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP
/PGECC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.

FEATURES source

Location/Qualifiers
1..1165
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U18528"
/ecotype="Columbia"
/note="This clone is in pUNI 51."

Gene CDS

1..1165
/gene="At4g18200"
1..1134
/gene="At4g18200"
/codon_start=1
/evidence=experimental
/product="putative protein"
/protein_id="AA020208.1"
/db_xref="GI:20465497"
/translation="MGISQVHYCNGDQNLNLDHEETESFVPTQKNCWLRVSI
YAIFFVPCPLAVLGRLYENGKSTVYVTLQLIGPVLILFRFFSRIPKPSKSTDI
NFSQSPFTLASVLCGLLVSAVYLSAVGLYLPVSTFSLIASQLAFAPFSYF
LNSQKFTPLIVNLSFLITVSSALLVNTDSENTNVSRVQYVIGICTIGASAGYGLV
LSLIQLFRKVFRTKTSALLVAVQSLVATCVVLIGLFGSEGWRTLPSEMRNYKLG
KVSYLTLASAAI FQWYTVGCVGLIFSSSVFSNSITAVGLPIPVVAVIVFHDKMD
ASKISFIILAIWGLFSLFVYQHYLDEKKLKTQKPVBEETQL"
1135..1165
/gene="At4g18200"

ORIGIN

Query Match 59.9%; Score 775; DB 8; Length 1165;
Best Local Similarity 83.8%; Pred. No. 1,2e+188;
Matches 877; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
QY 180 GGAATCGAATCTTCGTCGTAACCTCAATCGAAGAACTATAAGAAATGCGTTCGTATTC 239
Db 69 GGAATCGAATCTTCGTCGTAACCTCAATCGAAGAACTATAAGAAATGCGTTCGTATTC 128
QY 240 CATTTACGTCGTCCTTCTGTCCTTCCCAACACCTTTCTACAAATTTGGGAGAGATTGA 299
Db 129 CATATACGCAATCTTTGTCTCTCTGCAACACCTTGTCTACAGTTCTGGTAGACTGTA 188
QY 300 CTATGAAATGCTGGGAGAGATACATGATGGGAACACTTGTCCAACTAATCGGCTTCCC 359
Db 189 CTATGAAATGAGGAGGAAAGACATATGTGTGTAAACCTTCTTCAACTCATTTGGCTTCCC 248
QY 360 TGTTCGTGTTCTTCCTCGGCTTCTTTTCCAAACCAAAATCCAAACCAAGAGACAGA 419
Db 249 TGTACTGATCTCTTCCTCGGCTTCTTTTTCGAATCAGGCAACCAAAATCAACAGATACAAA 308
QY 420 TTTGAGAAAGTCTCTCTCTTCCATTCCTTGGATCAGTTTACATCGTTTACGGACTATT 479
Db 309 TTTGAGTCACTCCCTTCTTCCACCACTTGTGATCGGTTTACTTGTGACATGGACTGCT 368
QY 480 AGTGTCTGTAACCTCTTATATGTCCTCTGTTGGTTTACTATCTTACCAAGTTTCTACTTT 539
Db 369 AGTGTCTGTTATGCTTATTTGTCTGAGTGGTGGTTTCTTACTTACCAAGTCTCTACTTT 428
QY 540 TTCCCTCATCTTGGGCTTCAAAATTTGGGCTTCTCACTGCTTCTTCTCATATTTTCTAAACTC 599
Db 429 CTCCCTCATCTTGGGCTTCAAGTTTGGCTTCACTGCTTCTTCTCATATTTTCTTAACTC 488
QY 600 GCAGAGTTTCAACCTTTCATTTGGAATCTCTGTTTCTTCTTCTTCTTCTTCTTCTGCTCT 659
Db 489 GCAAAAGTTCACTCTCTTTGATAGTCAATTTCTTGTGTTCTTCTTCTTCTTCTTCTGCTCT 548
QY 660 CCTGCTGTCAACACTGATTCGGAACACACAGCAAAAGTGTCTAGAGTAAATATGTGAT 719
Db 549 CCTGCTGTCAACACTGATTCAGAAACACAAATGTATCTAGAGTAAATATGTGAT 608
QY 720 AGGGAATAATGTATACCTTGTGCTTCTGCTGGGATTTGGATTTGCTTCTTCTGCTGCTGATA 779
Db 609 TGGGTTCTATCTGTACCACTTGTGCTTCTGCTGGGATTTGGATTTGCTTCTGCTGATA 668
QY 780 ACTGATCTCAGGAAGGTTTAAAGAACAAACATCTCAACGGTCACTGACTTGGTGGC 839
Db 669 ACTGCTCTCAGGAAGGTTTCAAGAACATACATCTCTCAGCAGTCTCTGGACTTGGCCAA 728
QY 840 TTACCAATCTCTAGTTGCAAGCTGTGTTCTCTCATAGGACTTTTTCGGAAGGGGAGTG 899
Db 729 TTACCAAGTCTCTAGTTGCAAGCTGTGTTCTCTCATAGGACTTTTTCGGAAGGGGAGTG 788
QY 900 GAAAACTTTAAACAGTGAGATGGAACACTCAAACTGGGGAAAGTSCCATAGCTTATGAC 959
Db 789 GAGAACTCTGCCAAGTGAGATGAGAAACTCAAACTGGGGAAAGTGTATATATCTTGAAC 848
QY 960 TTTGGCTCTGATAGCTATTTCTGCGCAAGTCTACCAATTTGGCTGCTGGGAGTGAATCTT 1019
Db 849 TTTGGCTCTGATAGCTATTTCTGCGCAAGTCTACCAATTTGGCTGCTGGGAGTGAATCTT 908
QY 1020 TGAGTCATCTTCTGTTTCTCAATTTCCATTAACCTGCTGGGATTTGGCTTATAGTTCCAGT 1079
Db 909 CGAGTCTTCTGTTTCTCAATTTCCATTAACCTGCTGGGATTTGGCTTATAGTTCCAGT 968
QY 1080 TGTAGCAGTATGTTTTCATGATAAATGAACGCGTCAAAAGATCTTCTCCATCATTTT 1139
Db 969 TGTAGCAGTATGTTTTCATGATAAAGTGAAGTGAATCAAAAGATTTTCTCCATCATTTT 1028
QY 1140 AGCTATCTGGGATTCATTTTCTATGCTATCAGCACTACCTCGACGAAAAAGAGTTGAA 1199
Db 1029 AGCTATCTGGGCTTCTTCTTCTATGCTATCAGCACTACCTCGACGAAAAAGAGTTGAA 1088

3' UTR

Qy	1200	GACTAGCCACACAACTCTCTGTAGGAGA	1226
Db	1089	GACTTGCACAGAAAACCTGTGAGGA	1115
RESULT 9			
AY074546			
LOCUS	AY074546	1428 bp	mRNA linear
DEFINITION	Arabidopsis thaliana unknown protein (At4g18200)		mRNA, complete cds.
ACCESSION	AY074546		
VERSION	AY074546.1	GI:18491220	
KEYWORDS	FLI CDNA		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE			
AUTHORS	Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
TITLE	Arabidopsis Full Length cDNA Clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1428)		
AUTHORS	Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,B., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JAN-2002)		
COMMENT	Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K. The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.		
FEATURES			
source	Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.		
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.			
Location/Qualifiers	1..1428		
/organism="Arabidopsis thaliana"			
/mol_type="mRNA"			
/db_xref="taxon:3702"			
/chromosome="4"			
/clone="RAFL09-36-N15 (R18628)"			
/ecotype="Columbia"			
/note="This clone is in a modified pBluescript vector (FLC-1) as a BamHI/XhoI insert."			
gene	1..1428		

5'UTR	/gene="At4g18200"		
CDS	1..71		
	/gene="At4g18200"		
	72..1205		
	/gene="At4g18200"		
	/codon_start=1		
	/evidence=experimental		
	/product="unknown protein"		
	/protein_id="AA169512.1"		
	/db_xref="GI:18491221"		
	/translation="MGISQVHYCNGDQNLBLEANLLDHEETESFSPQTKNCKRWLRVSI YAIFFVFCQPLATVLRGLYYENGKSTYVVTLLQLIGFVPLILFRFSPRQAPSFV NFQSPSFETLASVYLCTGLVSAVLSAGLLYLFPVSFSLILASQLAFTAPSFV LNSQKFTPLIVNSLFLTLVSSALLVNTDSENTNSRVQYVIGFICTIGASAGIGLV LSLIQLLFRKVPFKHTSSAVLDLANYSQSLVATCVLLIGLPASGEWRILPEMRNYKLG KSYLILTLASAAIFMQVYTVGCVLGFESSVFSNSITAVGLPIVPVAVVAVIFPHDKMD ASKIFSILAIWGLFSFYQHYLDEKLLKTKQTKPVEETQTL"		
3'UTR	1206..1428		
misc_difference	/gene="At4g18200"		
	1412		
	/gene="At4g18200"		
	/note="not present in genomic sequence"		
ORIGIN			
Query Match	59.9%	Score 775;	DB 8; Length 1428;
Best Local Similarity	83.8%	Pred. No. 1.2e-188;	
Matches	877;	Conservative 0;	Mismatches 170; Indels 0; Gaps 0;
Qy	180	GGAAATCGAATCTTCGTCGGTACCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTC	239
Db	140	GGAAACTGAATCATTTTCAGTACCTCAACGAAAGCTGAAGAGGTGGCTCCGTCCTC	199
Qy	240	CATTACGTGTTCTTGTCTTCGTCGCAAGCACTTCTTACAAATTTGGGCAGAGTTTA	299
Db	200	CATATACGCAATCTTTGTCATCTTCTCCAAACCACTTGTACAGTTCTGGGTAGACTGA	259
Qy	300	CTATGAAATGGTGGGAAGAGTACATGGATGGGAACACTTGTCCAACTAAATCGGCTTCCC	359
Db	260	CTATGAAATGGAGGAAGACATATGTGTAACTCTTCACTCATTCATTCGCTTCCC	319
Qy	360	TGTTCTGTTCTCTTCGCTTCTTTTCCCAACCAAAAATCCCAACCAACAGAGAGA	419
Db	320	TGTAATGATTTCTTCGCTTCTTTTTCGAAATCAGGCAACCCAAATCAACAGATACAAA	379
Qy	420	TTTCAGAAAGTCTCTTCTTCCATCTTGGATCAGTTTACATCGTTACTGACTATT	479
Db	380	TTTCAGTCAGTCCCTTCTTCCACCCCTTGCATCGGTTTACTTGTGCATCGACTGCT	439
Qy	480	AGTGTCTGCTAACTCTTATATGCTCTCTGTGGTTTACTATATCTTACAGTTTCTACTTT	539
Db	440	AGTGTCTGCTTATGCTTATTTGTCGAGTTGGTTGCTTTACTTACAGTCTCTACTTT	499
Qy	540	CTCCCTCATCTTGGCTTCACAAATGGCTTCACGCTTCTTCTTCATATTTCTAAACTC	599
Db	500	CTCCCTCATCTTGGCTTCACAGTTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	559
Qy	600	GCAGAAATTCACACTTTCATTTGGAATTTCTGTTTCTCTTCTTCTTCTTCTTCTTCTTCTT	659
Db	560	GCAAAAGTTCACTCTTTGATAGTCAATTTCTTTGTTTCTCTTACGGTTTCTCTCGCCT	619
Qy	660	CCTCGTGTCAACACTGATTCGGAACACACAGCAAAAGTGTAGAGTAAATAATGTGAT	719
Db	620	CCTCGTGTCAACACTGATTCAGAAACACACAAATGTATCTAGAGTACAAATATGTAT	679
Qy	720	AGGGAATATATGACCAATTTGGTCTTCTGCTGGGATTTGGATTTGCTGCTATCCCTGGTACA	779
Db	680	TGGGTTTCATCTGTACCAATTTGGTCTTCTCGCTGGGATTTGGATTTGCTATCTCTGATACA	739
Qy	780	ACTGATCTCTGAGGAGGTTTAAAGAGCAAACTTCTCAACGCTCACTGACTTGGTCGC	839
Db	740	ACTGCTCTTCAGGAAAGTTTTCAGAAAGATACATCTCTCAGCAGTCTCTGCACTTGGCAA	799
Qy	840	TTACCAATCTCTAGTTCGAAGCTGTGTGGTTCTCATAGGACTTTTTCGCAAGCGGGAGTG	899


```

800 TTACCAAGTCTCTAGTTGCAACTTTGTGTGGTACTCATAGGACTGTGTTTGCAGTGAGAGTG 859
900 GAAAACTTTTAAACAAGTGAGATGAAAACTACAAACTGGGGAAGTGCCATACATTATGAC 959
860 GAGAACTCTGCCAAGTGAGATGAGAACTACAAACTGGGGAAGTGCTATATATCTTGAC 919
960 TTTGGGCTCGATAGCTATTCTTCTGGCAAGTCTACACCAATTGGCGTCTGGGACTGATCTT 1019
920 TTTGGGCTCAGCAGCTATTCTTCTGGCAAGTATACACTGTTGGTGTGGGATTAATCTT 979
1020 TGAGTCATCTTCTGTCTCTCCAAATTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCT 1079
980 CGAGTCTTCTTCTGTCTCTCCAAATTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCT 1039
1080 TGFAGCAGTGATTGTTTCTCCATGATAAAATGAACGGCTCAACAGATCTTCTCCATCAATTT 1139
1040 TGFAGCAGTGATTGTTTCTCCATGATAAAGATGATGATGATGATGATGATGATGATGATGAT 1099
1140 AGCTATCTGGGATTCATTTCTATCTATCAGCACTACCTCGACGAAAGAGTTGAA 1199
1100 AGCTATCTGGGCTTCTCTTCTATCTATCAGCACTACCTCGACGAAAGAGTTGAA 1159
1200 GACTAGCCACAAAGTCTGTAGGAGA 1226
1160 GACTTCCGACAAACCTGTTGAGGA 1186

RESULT 10
AF370622 1106 bp mRNA linear PLN 30-APR-2001
LOCUS Arabidopsis thaliana putative protein (T9A21.60) mRNA, complete
DEFINITION cds.
ACCESSION AF370622
VERSION AF370622.1 GI:13877726
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1106)
Lam,B., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R.,
Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H.,
Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (17-APR-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arab@sequence.stanford.edu
This clone was isolated by RT-PCR.
FEATURES
source
    1..1106
        /organism="Arabidopsis thaliana"
        /mol_type="mRNA"
        /db_xref="taxon:3702"
        /chromosome="IV"
        /clone="C531"
        /ecotype="Columbia"
    1..1106
        /gene="T9A21.60"
        657..1106
        /gene="T9A21.60"
        /codon_start=1
        /product="putative protein"
        /protein_id="AAK43941.1"
        /db_xref="GI:1387772"
        /translation="MDMIIYVSLVASCVWGLFASSEWKTLSSEMDNYKHGKVSYIM
        NLVWATVWQVFSIGTGLIFELSLFSNALSVGLPVVPIILAVLIFHDKNGLKVIS
        MILAIWGFYSYVYQQYLDLDDLNKLNHEITTESPDPEABESTWQSK"

gene
CDS
ORIGIN
```


	roids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1.
AUTHORS	Glaizebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.
TITLE	Plant genes, the expression of which are altered by pathogen infection
JOURNAL	Patent: WO 0222675-A 632 21-MAR-2002; Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl, Jeffrey L. (US) ; Eulgem, Thomas (US)
FEATURES	Location/Qualifiers source 1..1047 /organism="Arabidopsis thaliana" /mol_type="unassigned DNA" /db_xref="taxon:3702"
ORIGIN	
	Query Match 31.7%; Score 410; DB 6; Length 1047; Best Local Similarity 65.0%; Pred. No. 1.4e-94; Matches 624; Conservative 0; Mismatches 330; Indels 6; Gaps 1;
Qy	293 GAGTTTTACTATGAAATGGTGGAAGAGTACATGGATGGGAACAACCTGTGTCCAACTAATCG 352
Dd	71 GAGTATACTATGACAA CGGAGGAACAAGTAATGGTAGCAACGGTAGTTCAACTTGTG 130
Qy	353 GCTTCCCTGTTCTGTCTCTCTCCGCTCTTTTTCCCAAACCAAATAATCCCAAACAACAG 412
Dd	131 GCTTTCTGTGCTACTTCCATATTATATATCTTGTCATTTAAAACACATGCACAACTGATA 190
Qy	413 AAGCAGATTTCAGAAAGTTCTCTTCCTTACCATTCTTGGNATCAGTTTACATCGTTACTG 472
Dd	191 GAGATGGAAGAAGAACCTCACT-----AGGAACCGGTGATTGGTTTACGTAGTGGTGG 244
Qy	473 GACTATTAGTCTGTCTAACTCTTATATGTCTCTGTGGTTTACTATATCTTACACAGTTT 532
Dd	245 GACTCTGTGAGGAGCAGATTGCTATCTGTACTCATTTGGNACTTCTTATCTACCCTGTT 304
Qy	533 CTACTTTTCCCTCATCTTGGCCCTCAAAATTGGCCTTCACTGCCCTTCTTCTCAATATTTC 592
Dd	305 CTACCTATTCCCTGATCTGTGCATCTCAGTTTAGCCTTCAATGCTTCTTCTCTTATTTC 364
Qy	593 TAAACTCGCAGAACTTCACACCTTTCAATGTGAATTCCTGTGTTCTCTGTACTATTTCCT 652
Dd	365 TTAACTCACAAAAACCTTACCCCTATCAATTTTAAATCTCTTTTCTCCTCTAACTATATCT 424
Qy	653 CTGCCCCCTCTGGTCAACACTGATTTCGGAACACACAGAAAAAGTGTCTAGAGTAAAT 712
Dd	425 CCACCTACTTGCAATTCATATGAGGAGACAGACTCCACAAAAGTTTACAAAGGAGAGT 484
Qy	713 ATGTGATAGGGATTAATATGTACCAATGGTGCTTCTCTGCGGATTTGGATTGCTGCTATCCC 772
Dd	485 ATGTCAAAAGGTTTCATATGCACCGTTGCTGCGTCTGCTGGTTATGGTCTAGTCTTATCCC 544
Qy	773 TGGTACAACTGATCCTCAGGAGGTTTTAAAGAAGCAAAACATTCCTCAACGGTCACTGACT 832
Dd	545 TACAACAGCTAGCCCTTTCTAAAGTCCTTAAAGAAGCAAAATTTCTCAGAAGTTATGGATA 604
Qy	833 TGGTGCGCTTACCAATCTCTAGTTCGACGCTGTGGTTCTCATAGGACTTTTTCGACAGCG 892
Dd	605 TGATAATCTACGTGAGTCTAGTGCCCAAGTTGTGTTAGCGGTGGGGCGCTTTTGTCTAGCA 664
Qy	893 GGGAGTGGAAAACTTTTAAACAAGTGAGATGGAAAACTACAACTGGGAAAAAGTGCCATACG 952
Dd	665 GTGAGTGGAAAACTTTTGAGCAGTGAATGGATACTACAAACATGGGAAGGTATCCTTACA 724
Qy	953 TTATGACTTTGGCCTCGATAGCTATTTCCTGGCAAGTCTACACAAATTGGCGTCTGTGGAC 1012
Dd	725 TTATGAACCTAGTGTGGACAGCTGTATCTCTGGCAGAGTAATCTCCATCGGTGGCACAGAC 784
Qy	1013 TGATCTTTGAGTCATCTTCTGTGTTCTTCCAATTCATAACTGCTGTGGGATTGGCCTATAG 1072
Dd	785 TGATCTTCGGAGCTCTCTCTCTCTATCTCTCAATGAATGAAGCGTTTGGGACTCCCACTGG 844
Qy	1073 TTCAGTTGTAGCAGTGAATTTGTTTTCATGATAAAATCAACCGCTCAAGAAGATCTTCTCCA 1132

Qy	773	TGGTACAACTGATCCTCAGGAGGTTTAAAGAGCAAAACATTTCTCAACGGTCACGTGACT	832
Db	563	TACAACAGCTAGCCCTTTCTAAAGAGTCTTAAGAGCAAAATTTCTCAGAAGTTATGGATA	622
Qy	833	TGGTCGCTTACCAATCTCTAGTTGCAAGCTGTGTCTCTCATAGGACTTTTCGCAAGCG	892
Db	623	TGATAATCTACGTGATCTAGTGCCAGTTGTGTAGCGTGGTGGGCTTTTGTCTAGCA	682
Qy	893	GGGAGTGGAAAACTTTTAAACAAGTGAGATGGAAAACTTACAAACTGGGGGAAAGTGCCATACG	952
Db	683	GTGAGTGGAAAACTTTTGAGCAGTGAAATGGATAACTTACAAACATGGGAAGGTATCCTACA	742
Qy	953	TTATGACTTTGGCCTCGATAGCTATTTCTTGGCAAGTCTACACCATTTGGCGTGGTGGGAC	1012
Db	743	TTATGAACCTAGTGTGGACAGCTGTTACCTGGCAGGTATTTCTCCATCGGTGGCACAGGAC	802
Qy	1013	TGATCTTTGAGTCATCTTCTGTCTCTCCAAATTTCCATAACTGTGTGGGATTGCCCTATAG	1072
Db	803	TGATCTTCGAGCTCTCTCTATTTCTCAATGCAATAAGCGTTTGGGACTCCCAAGTGG	862
Qy	1073	TTCCAGTTGTAGCAGTGATTTGTTTCCATGATAAAAATGAACGCGTCAAAAGATCTTCTCCA	1132
Db	863	TTCCATCTTTGGCTGTAAATCAITTTCCATGACAAAATGAATGGGTTAAAGGTGATTTCTA	922
Qy	1133	TCATTTTAGCTATCTGGGGAATTCATTTCTATCAGACACTACCTCGACGAAAAAGA	1192
Db	923	TGATCCTAGCTATTTGGGGTTTCACTTCTCTATGTCTACCAACAATATCTTGATGACAAA	982
Qy	1193	AGTTGAAGACTAGCCACACAAGTCCGTAGGAGATCCTCATCTACTACCTGCTGAGGAG	1252
Db	983	ACTTGAAGAAAAATCATGAATCACAAACAGATCCCTGACCCACCAGAGCAAGAAG	1042

Search completed: November 1, 2004, 15:44:38
Job time : 5735.91 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 06:20:04 ; Search time 593.039 seconds
(without alignments)
9900.026 Million cell updates/sec

Title: US-09-913-767-3

Perfect score: 1145

Sequence: 1 ctgttctatgttggatgtg.....caagtttaggcaaatgaccca 1145

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09E_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	354	30.9	1294	16	US-10-424-599-2524
2	318.4	27.8	3387	9	US-09-938-842A-2212
3	318.4	27.8	3387	11	US-09-938-842A-2212
4	281.8	24.6	1513	17	US-10-437-963-82061
5	280.4	24.5	2214	16	US-10-424-599-102519
6	243.2	21.2	83698	9	US-10-416-898-9
7	217	19.0	997	16	US-10-424-599-75161
8	197.2	17.2	2381	17	US-10-437-963-43860
9	168.6	14.7	728	17	US-10-767-701-9000
10	163	14.2	747	17	US-10-767-701-6806
11	154.6	13.5	592	9	US-09-770-152-44
12	130.6	11.4	2605	17	US-10-437-963-83900

13	128.4	11.2	792	16	US-10-424-599-9213	Sequence 9213, Ap
14	127	11.1	1779	16	US-10-425-114-28473	Sequence 28473, A
15	125.4	11.0	1386	16	US-10-425-114-12485	Sequence 12485, A
16	125	10.9	1432	17	US-10-437-963-1897	Sequence 1897, Ap
17	111.8	9.8	522	16	US-10-424-599-114931	Sequence 114931, A
18	105.2	9.2	327	16	US-10-424-599-121152	Sequence 121152, A
19	104.4	9.1	332	11	US-09-732-627A-1196	Sequence 1196, Ap
20	97	8.5	1146	16	US-10-425-114-10516	Sequence 10516, A
21	97	8.5	1233	16	US-10-424-599-30274	Sequence 30274, A
22	87.4	7.6	1373	16	US-10-424-599-31364	Sequence 31364, A
23	86.6	7.6	650	17	US-10-767-701-4045	Sequence 4045, Ap
24	84	7.3	1333	16	US-10-425-114-10574	Sequence 10574, A
25	83.4	7.3	671	17	US-10-767-701-5986	Sequence 5986, Ap
26	78.4	6.8	1355	16	US-10-424-599-31363	Sequence 31363, A
27	76.2	6.7	774	9	US-09-770-445-895	Sequence 895, App
28	75	6.6	308	16	US-10-424-599-121805	Sequence 121805, A
29	53.2	4.6	1417	16	US-10-424-599-97043	Sequence 97043, A
30	52.8	4.6	386	9	US-09-770-791-85	Sequence 85, Appl
31	52.8	4.6	716	16	US-10-424-599-61310	Sequence 61310, A
32	49	4.3	13573	15	US-10-311-455-1842	Sequence 1842, Ap
33	47.4	4.1	613	17	US-10-767-701-4026	Sequence 4026, Ap
34	45	3.9	3090	16	US-10-654-416-5	Sequence 5, Appli
35	44.6	3.9	17213	15	US-10-311-455-1456	Sequence 1456, Ap
36	44	3.8	15373	15	US-10-311-455-440	Sequence 440, App
37	43	3.8	6145	16	US-10-221-714A-269	Sequence 269, App
38	42.6	3.7	3081	16	US-10-654-416-9	Sequence 9, Appli
39	42.6	3.7	46732	17	US-10-741-601-5752	Sequence 5752, Ap
40	42.4	3.7	492	16	US-10-424-599-92714	Sequence 92714, A
41	42.2	3.7	631	17	US-10-767-701-5979	Sequence 5979, Ap
42	42	3.7	402	16	US-10-131-837-8426	Sequence 8426, Ap
43	42	3.7	3673778	15	US-10-312-841-1	Sequence 1, Appli
44	41.6	3.6	476	10	US-09-814-353-4700	Sequence 4700, Ap
45	41.6	3.6	476	10	US-09-814-353-10999	Sequence 10999, A

ALIGNMENTS

RESULT 1
US-10-424-599-2524
; Sequence 2524, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 2524
; LENGTH: 1294
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_102285C.1
US-10-424-599-2524

Query Match	30.9%	Score 354;	DB 16;	Length 1294;
Best Local Similarity	59.6%	Pred. No. 2.9e-89;		
Matches 619;	Conservative 0;	Mismatches 410;	Indels 9;	Gaps 1;
QY	33	GAATATTTTCAACAGAGAAGTCAACAAGTACTCTTGGAGGTTAAGAGTGCTCTCTA	92	
Db	94	GAATGAATCAACAATGAGTAAGAAAAAAGATATTACCGTTGGCTCAGAAATATCCATCCA	153	
QY	93	TGTCACTCTCTCTTAGCTGGAGAGACAATAGGCACTCTCTTAGTACACTTATACGA	152	
Db	154	TTCTTCACTTGTGTGTGGTGTGCGGATCAGCAGCAATCTCTCTTGAAGATTGTACTATGA	213	

```
Qy 153 AAGAGCGGTAAAGCACATGGCTCGAAACCTTGGTTACAGTTTCTAGGTTTCCCTTAAAC 212
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 214 AAGAGGTGGAAGAAAGCAAGTGGATGGGAACACTTGTTCAACTTGTGTTTCCCTATTCA 273
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 213 CCTTCCTTGTCTATTATTACTTAAAGCGCTGAGCCGTCGAAGACTAAACCAATTACCAAAA 272
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 274 GCTTCCTTTTCATTTATTTCAGCATCAAAAATCTCACCAAAATAGTAGCAATTCATCC 333
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 273 AACTACTTCTCTCTTTGACACTATCTTTTAGTGTATATGGAAGCTGGCTGGCTTGTGTC 332
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 334 AAAACAATCATCTGCTTCAATACTAGCAATTTATCTATGCTCTCAATTTGGCTACTTTGGC 393
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 333 TGGCAATGTATTTGTACTAATTTGGCTACTTTTACCTTCTGCTCTCAACTTCTCTTT 392
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 394 ATTAGATTGTATTATTATTAGTTGAGTATGGAGCCTTCTGTCTCTACTATTCACT 453
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 393 GATCTCTGCTCGCAATTGGCTTTTAAACGCGTCTTCTTCTTACTTCTTAACTCACAAA 452
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 454 CATTTGCTCATCCCAATTTGGCTTTCAATGCTTTCTTCTCTACTTCTCTCACTCAAA 513
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 453 AATCACACATTTATFACATCAATTCATCTTTTGTCTTTAAACATATCTTCTACACTTCTGT 512
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 514 GTTCACACCTTACATAATCAACTCTCTAGTCTCTCTCACCATTCTTCAACCCCTCTGT 573
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 513 TATCCACATGAACCAAGATC-----TCCCTCTCTACTTCAAGTCCGAGCAA 563
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 574 GTTTCAAAATGAGTCATCATCAGATGATGATGATTCAGATTCACAAAATCTCCAA 633
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 564 GTCCAAGTATGTGATTTGATACATCTGCGCGTGGTAGCTCAGCTGCTGTTATTTCTCTGGT 623
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 634 GAAAAGTATGTGATTTGATTCATATGCACAGTTGGTGTGCACTCGAGGGTATGACTATG 693
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 624 GCTTTCTTTAAACAGATTACGGCTTCGAAAAGATTCTAAAGAAATFACATTTCAAGGCTAT 683
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 694 GCTTTCCCTCACACAGCTTGTGTTCAAGAAAGTTCATAAAAAGAGAAAACATTCAAAAGTGT 753
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 684 TTTAGACATGCCACATATCGCTATGTTAGTACTTGTGTAGTGTGTTAGTACTTTT 743
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 754 CTTGGATATGATATTATACATTTCCCTTGTGGTACCCCTTGTACTACCTTAGTGGACTTT 813
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 744 TGGAGTGGTGGTGGAAAAAGCTGAGTACAGAAAATGGAAGATTTCAACTAGGGAAGAAG 803
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 814 TGCTAGTGGAGTGGAGTGTGTTGAAGAATGAATGAAGAGTATGATGGGGAGGC 873
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 804 CTCATACATTTTGAATAAACATCGGTTTCAACGATATCATGCAAGCTGTTTGTATGGAAG 863
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 874 TTCATATTGCTGAACCTCACTTTTACAGCCATATCTTTGGCAAGTCTTTCAACATTTGGTTG 933
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 864 TGTGGTTTGTATATCGAAGTTTCATCGCTTTTCCAAATGTCATAGCACTCTTTTGT 923
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 934 TTTGGCGCATATTAGTGAGGTTTCTCCCTCTTCTTAATGCCATAAGCGCTTTGGGAGT 993
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 924 ACCAGTTTGTCTGTTCTTGTGTTGTTCTTCTCGGTGATGAGTGGAAATCAAGTT 983
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 994 GCCTATTCTTCAATGCTGGCAGTGTGTTCTTTCATGACAAAATGGATGCAATTAAGG 1053
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 984 GGTTCGAATGTTTGGCCATCTGGGAAATTTGTTTCTTATGGTTATACGCAATATGCA 1043
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1054 TATCTCTATGTTGTAGCTATTGCGGCAATGTATCATATGTATCAACAGTACTTGA 1113
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1044 TGATAGAACCCAGAGA 1061
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1114 TGATACAAAATCCGAAAA 1131
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 2

```
US-09-938-842A-2212
; Sequence 2212, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
```

```
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2212
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-2212
```

```
Query Match 27.8%; Score 318.4; DB 9; Length 3387;
Best Local Similarity 57.8%; Pred. No. 6.4e-79;
Matches 593; Conservative 0; Mismatches 421; Indels 12; Gaps 1;
```

```
Qy 28 GAGGGGAAATTTTCAACAGAGAGAGAAAGTCACAAGTACTCTTGGAGGTTAAAGAGTGTCT 87
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 88 CTCTATGTCACTCTCTCTTTAGCTGGAGAGACAATAGCCACTCTCTTTAGGTAGACTTTTAC 147
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 148 TACGAAAAGCGGTAAAGACACATGGCTCGAAACCTTGGTTCAGCTTGTAGGTTTCCT 207
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2443 TATGAAAATGGAGGGAAGACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCT 2502
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 208 TTAACCCCTCTCTGTCTATTATTACTTAAAGCTGAGCGCTCCAAGACTTAAACCATTAAC 267
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2503 GTACTGATCTCTCCGCTTCTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAAT 2562
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 268 AAAAAAACTACTTCTCTCTTTGACACTATCTTTAGTGTATATTTGGAATTTGGCTTGGCTT 327
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2563 TTCAGTCAGTCCCCCTTCCCTTCCACCCTTGCATCGGTTTACTTGTGCATCGACTGCTA 2622
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 328 GTTGTGGAATGTATTTTGTACTCATTTGGGCTACTTTTACCTTCTCTGCTCAACTTTC 387
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2623 GTGTCTCTTATGCTTATTTGTCTGCAATTTGTTTGTCTTCTCTTACCTTACCACTTCT 2682
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 388 TCTTTGATCTCTGCTCGCAATTTGGCTTTTAAAGCCGCTTCTCTTACTTCTCTTAAACTCA 447
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2683 TCCTCATCTTGGGCTCACAGTTGGCCTTCACTGCTTTTCTCATATTTCTTAACTCG 2742
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 448 CAAAAATACACCAATTTACTCAATTCATCTTGTCTCTTTAACCATATCTTCTACACTT 507
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2743 CAAAAAGTCACTCTCTTGATAGTCAATTTCTTGTCTTCTCTTACGTTTCTCTGCGCTC 2802
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 508 CTGTGTTATCCAAATGAACAGATCTCCCTCTTCTTACTTCAAGTCCGAGCCCAAGTCC 567
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2803 CTCGTGTCAACACTGATTCAGAA-----AACACAACAATGATCTAGAGTA 2850
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 568 AAGTATGTGATTTGATACATCTGCGGCTCGGTAGCTCAGCTGGTTATTTCTCTGCTGCTT 627
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2851 CAATATGTGATTTGGTTTCATCTGTACCATTTGGTCTTCCGCTGGGATTTGGATTGTA 2910
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 628 TCTTTAAACAGATTACGGTTTCGAAAAGATTCTAAGAAATACACATTTCAAGGCTATTTTA 687
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2911 TCTCTGATACAACTGCTCTTTCAGGAAAGTTTTCAGGAAGCATATACCTCTCAGCAGTCTG 2970
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 688 GACATGCCACATATCCGTCTATGTTAGTACTTGTGTAGTGTGTTGGTGGAGCTTTTGA 747
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2971 GACTTGGCCAAATTACCAGTCTCTAGTTGTGCAACTTGTGTGTACTCATAGGACTGTTTGA 3030
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 748 AGTGGTGGTGGAAAAAGCTGAGTACAGAAATGGAAGATTTTCAACTAGGAAAAAGCTCA 807
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 3031 AGTGGAGAGTGAGAACTCTGCCAAGTAGAGAGAGAACTACAAACTGGGAAAGTGTC 3090
Qy 808 TACATTTTGATAAACAATCGGTTCAACGATATCATGGCAAGCTTGTGTGATTCGAAGTGTT 867
Db 3091 TATACTTGACCTTGGCCCTCAGCAGCTATTTCTGCAAGATATACACTGTTGGTGTGG 3150
Qy 868 GGTGATGATTCGAAGTTTCATCGCTTTTTCGAATGTCATGAAGCACTCTTTGTTTACCA 927
Db 3151 GGAATTAATCTCGAGTCTTCTCTGTGTCTCCTCAATTCATAAATCTGCTGGGACTGCT 3210
Qy 928 GTTGTGCGCTTCTGCTGTCTCTTCTCGGTGATGAGATGAGATGAGATGAGATGAGAT 987
Db 3211 ATAGTTCAGTTGTAGCAGTAGATTTTCCATGATGAAGATGATGATGATGATGATGAT 3270
Qy 988 GCAATGTTTGGCCATCTGGGATTTGTTCTTATGTTTATGAGTATGATGATGATGATGAT 1047
Db 3271 TCCATCATTTTAGTATCTGGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3330
Qy 1048 AGAAG 1053
Db 3331 AAGAAG 3336

RESULT 3
US-09-938-842A-2212
; Sequence 2212, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krepe, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2212
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2212

Query Match 27.8%; Score 318.4; DB 11; Length 3387;
Best Local Similarity 57.8%; Pred. No. 6.4e-79;
Matches 593; Conservative 0; Mismatches 421; Indels 12; Gaps 1;

Qy 28 GAAGGAAATTTTCAACGAGAGAGAGTCAAGTACTCTTGGAGGTTAAGAGTGCT 87
Db 2323 GAAACTGAATCATTTTTCAGTACTCTCAACGAAAGACTGTAAGAGTGCTCGTCTCTCC 2382
Qy 88 CTCTATGCTACTCTCTTCTAGCTGGAGAGCAATAGCCACTCTCTTAGGTAGACTTTAC 147
Db 2393 ATATACGCAATCTTTGTCTATCTTCTGCAACCACTTGTCTACAGTTCTCGGTAGACTGTAC 2442
Qy 148 TACGAAAAGGCGGTAAAGACATCGGCTCGAAACCTTGGTTTCAGCTTGTAGGGTTTCTCT 207
Db 2443 TATGAATATGAGGAAAGACATATGTGTTAACAATCTTCTCAACTCATTTGGCTTCTCT 2502
Qy 208 TTAAACCTTCTTCTGCTATTTATTTAAGCTCGAGCCGTCAAGACTTAAACCAATACC 267
Db 2503 GTACTGATCTCTTCTCGCTTCTTCTTCTGCAATCAGGCAACCAATCAACAGATACAAAT 2562
Qy 258 AAAAAAATCTCTCTCTCTTCTGACACTATCTTTAGTGTATTTATTTGGACTTGGCTTCT 327
Db 2563 TTCACTCAGTCCCTTCTCTTCAACCCCTTGCATCGGTTTACTTGTGCACTGGACTGCTA 2622

Qy 328 GTTGTGACATGTAATTTTGTACTCATTTGGGCTACTTTACCTTCTCTCTCAACTTTC 387
Db 2623 GTGCTGCTATGCTTATTTGTCTGCAGTTGGTGTGCTTACTTACCAGTCTCTACTTTC 2682
Qy 388 TCTTTGATCTCTGCGTCGCAATTTGGCTTTTAAAGCCGCTTCTCTTACTTCTCTTAAACTCA 447
Db 2683 TCCCTCATCTTGGCCCTCAGTTGGCTTCTACTGCTTTTCTCATATTTCTCTTAACTCG 2742
Qy 448 CAAAAAATCACACCAATTTATATCAATTCACCTTGTGTTCTCTTAAACATATCTCTACACTT 507
Db 2743 CAAAAGTTCACTCTTGTATAGTCAATTTCTTGTCTCTTACGGTTCTCTGCCCC 2802
Qy 508 CTTGTTATCCAAATCAACCAAGTCTCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 567
Db 2803 CTCGGTCAACACTGATTCAGAA-----NACACAAATATCTAGAGTA 2850
Qy 568 AAGTATGATTTGGATATACATCTGCGCGGTGCTAGCTCAGCTGGTTATTTCTCTGTGCTT 627
Db 2851 CAATATGATTTGGGTTTCACTGTACCATTTGGTGTCTCCGCTGGGATTTGGATTTGTTACTA 2910
Qy 628 TCTTTAACAGATTACGCGTTGAAAGATTTCTAAGAAATACACATTCAGGCTATTTTA 687
Db 2911 TCTCTGATACACTGCTCTTTCAGAAAGTTTTCAGAAAGATACATCTCTCAGCAGTCTCTG 2970
Qy 688 GACATGGCCACATATCCGTCTATGTTAGTACTTCTTGTAGTTTGTGTAGGACTTTTGTGGA 747
Db 2971 GACTTGGCCAAATACCAGTCTCTAGTTGCAACTTGTGTGTTACTCATAGGACTTGTGGA 3030
Qy 748 AGTGTGGGTGGAAAAAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGGGAAGAGTCA 807
Db 3031 AGTGAGAGTGGAGAACTCTGCAAGTGAAGATGAGAACTCAAACTGGGAAAGTGTCA 3090
Qy 808 TACATTTTGATAAACAATCGGTTCAACGATATCATGCGCAAGCTTGTGTTGAGAGTGT 867
Db 3091 TATATCTTGACTTTGGCCCTCAGCAGCTATTTCTGGAAGTATACACTGTTGTTGTGTG 3150
Qy 868 GGTTTGATTTATCGAAGTTTTCATCGCTTTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTC 927
Db 3151 GGAATTAATCTTCGAGTCTTCTCTGTTCTTCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 3210
Qy 928 GTTGTGCTGTTCTTGTGTTGTTCTTCTTCTGTTGATGAGATGAGTGGAAATCAAGTTGTT 987
Db 3211 ATAGTTCCAGTTGTAGCAGTATAGTTTTCATGATGAAGATGATGATGATGATGATGAT 3270
Qy 988 GCAATGTTTGGCCATCTGGGATTTGTTTCTTATGTTTATGTTTATGATGATGATGATGAT 1047
Db 3271 TCCATCATTTTAGTATCTGGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3330
Qy 1048 AGAAG 1053
Db 3331 AAGAAG 3336

RESULT 4

US-10-437-963-82061/c
; Sequence 82061, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 82061
; LENGTH: 1513
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81525C.1
US-10-437-963-82061

Query Match 24.6%; Score 281.8; DB 17; Length 1513;
Best Local Similarity 57.0%; Pred. No. 9.4e-69;
Matches 564; Conservative 0; Mismatches 407; Indels 18; Gaps 2;

```
Qy 70 TGGAGGTTAAGAGTGTCTCTATGTACATCTCTCTAGCTGAGAGACAATAGCCACT 129
Db 1258 TGGTGGCTGATGGTGGAGTTAAATGTTCTCTCATTTGCCGTGACAGCGCATCTACA 1199

Qy 130 CTCTTAGTAGACTTTACTACGAAAAAGCGGTAAAGACACATGGCTCGAAACCTTGGTT 189
Db 1198 CTCTTAGGAGGTTCTACTACAATCAAGCGGCAATAGCAAGTGGATGTCCACATTCGTC 1139

Qy 190 CAGCTTGAGGTTTCCTTTAAACCTTCCTTGTCTATTATTACTTAAAGCCGTGACCGTCC 249
Db 1138 CAAACTGCTGGCTTCCGATTTGTTGATTCGCCCTATTCTTTCCATTCAAAGACATCT 1079

Qy 250 AAGACTAAACCATACCAAAAAA-----CTACTTCTTCTCTTGCACACTATCTTA 303
Db 1078 TCTACAAAAAGTCACTAGTAGTCTGCGCCCTACAATTTCTATCCCAAAATTTACTCTG 1019

Qy 304 GTGTATATTGACTTGGCTTCTGTTGCTGGACATTTGATTTTGTACTCAATTTGGGCTA 363
Db 1018 ATATATGTTGCTTGGGCTCATCATTTGCTGCAGACGACTTGATTTCTATGGGCTA 959

Qy 364 CTCTACCTTCCTGTCTCAACTTTCTCTTTGATCTCTGCTGCGCAATTTGGCTTTTAAAGCC 423
Db 958 CTATATCTTCGGGTCTCAACATATTGCTCATCTGTGTAGTCAGCTTGCCTTCAATGCT 899

Qy 424 GTCTTCTTACTTCTTAACTCAAAAAATCACACATTTATATCTCAATTTCACTTCTTGT 483
Db 998 GTCTTCTCATATTTCTCAATGCTCAAAAAATTCACCCCTCTGATTTTCAATTCGGTAGTC 839

Qy 484 CTCTTAAACCATATCTTCTACACTTCTTGTATCAACATGAACCAAGATCTCCCTCTTCT 543
Db 838 CTCTTAGCTTTCTGCTTCACTCTTGGAGTGTGATGAAGATTTCTAGGGAATCTACTAGT 779

Qy 544 ACTTCAAAGTCGAGCAAGTCAAGTATGTGATGTGATATATATATAGTATGATGCGGTGAGC 603
Db 778 ATATCA-----CATGGGAAGTACATTTTGGGTTTCTGTGTGACACTAGGGGCA 731

Qy 604 TCAGCTGGTTATCTCTGGTCTTCTTTTAAACATTTACGATTTACGGTTTCAAGAGATTTAAAG 663
Db 730 TCAGCTACATCTCGCTCATTTCTCTCCCTGATGCAAGTCAATTTGAGAGGTTTATTAAG 671

Qy 664 AAATACACATTTCAAGGCTATTTAGACATGCCACATCTCGTCTATGGTAGTCTTGT 723
Db 670 AGGAGAGCTTCTCAGTTGTGTGAAATGATGAGATATATACAGTCTCTCGTGGCAATG 611

Qy 724 GTAGTTGTGAGACTTTTGGAAAGTGGTGGGTTGGAAAGAGTGAATGAGTACAGAAATGAA 783
Db 610 GCTTCTCTTGTGGTTATTTGCAAGTGGTGAATGGATGACTTTTCAAGGAGAGATGCAT 551

Qy 784 GAGTTTCAACTAGGGAAGCTCATACATTTTGTGATAACATCGGTTCAAGCATATCATGG 843
Db 550 GCATTTCCAGTCTGGGAAGCTCTCATATGTAATGACACTGCTGTGGAGCGGCTATATCTGG 491

Qy 844 CAAGCTTGTGTTGATTTGGAAGTGTGGTTGATTTATCGAAGTTTTCATCGCTTTTTCCTCAAT 903
Db 490 CAGGTAGCATCAGTTGGAGTGTGGGATTTGATCTTTGTGGTGTCTATCGCTGTTTCAAT 431

Qy 904 GTCATAGACATCTTTGTTTACAGTTGTGCTGCTGTTCTTGTGTTGCTTCTTCCGCTGAT 963
Db 430 GTGATAAGCACCCCTAGCTCTACCCATCATTTCTCTGTTTGTGTGATTTTCTTTCATGAC 371

Qy 964 GAGATGAGTGAATCAAGTTGGTGAATGTTTGTGGCCATCTGGGATTTGTTTCTTAT 1023
```

```
Db 370 AAGATGATGGAGTAAGATTATAGCTATGCTATGATGCCCATTTGGGATTTATGTCATAT 311
Qy 1024 GGTATCAGCATTTATGTCAATGATAGAAA 1052
Db 310 GGCCACCAATTATATGTTGATGGCAAGAA 282

RESULT 5
US-10-424-599-102519
; Sequence 102519, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 102519
; LENGTH: 2214
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63592C.1
US-10-424-599-102519
```

```
Query Match 24.5%; Score 280.4; DB 16; Length 2214;
Best Local Similarity 57.0%; Pred. No. 2.9e-66;
Matches 583; Conservative 0; Mismatches 416; Indels 23; Gaps 3;

Qy 70 TGGAGGTTAAGAGTGTCTCTATGTCACTCTCTCTTAGCTGGAGAGACAATAGCCACT 129
Db 408 TGGTGGTGTCTTGTGACACTCAGCATAGCCCTTCTTATAGTTGGCAATCTGCTGCTGTT 467

Qy 130 CTCTTAGGTAGACTTTACTACGAAAAAGCGGTAAAGACATGGCTCGAAACCTTGGTT 189
Db 468 ATCCTTGGAAAGATTTTATTATGATCAGGTTGAAATAGTAAATGGATGGCTACTCTAGTT 527

Qy 190 CAGTTGTAGGTTTCTTTTAAACCTTCTCTGCTATTTACTTAAAGCTGAGCCGCTCC 249
Db 528 CAAAGCTGCTGCTCCCGAT--CTTGTCTCATTCATTTTACAATTCCTTCACCTCCAG 585

Qy 250 AAGACTAAAAACCATTAACAAAAAACTCTTCTCTTGTGACACTATCTTTAGTGTAT 309
Db 586 AGGCTTCAACTCTGCTTCACTCCCATCAAAATATTCTTTGATATATTTGGTCT-- 643

Qy 310 ATTGGACTTGGCTTGGCTTGGTGGACATTTGATTTTGTACTCATTTGGGCTACTTTAC 369
Db 644 -----TTGGAGTCTTAAATTTGCTGTGACAAATATGATGTACTCCACTGGACTCTTATAC 696

Qy 370 CTTCCTGTCTCAACTTCTCTTTGATCTCTGCGTGGCAATTTGGCTTTTAAAGCGCTTTC 429
Db 697 CTCTCGGCTTCTACCTTATTCGCTGATTTGTGCATCAGTAGCTTTTAAATGCAAGTTTC 756

Qy 430 TCTTACTTCTTAACTCACAATAAAATCACAACCATTTTATCTCAATTCACATTTGTTCTTTA 489
Db 757 TCATATTTTATCAATTTCTCAAAAGTTTCACTGCTTGTATTAATACTACAGTGGTTCTC 816

Qy 490 ACCATATCTTCTACACTTCTTGTGTTATCCAAATGAAACAGAAATCTCCTCTTCTACTTCA 549
Db 817 ACTTATCTGCTGCACTCTCTTGTGTTTAAAGGAAACAGACAGATGAACCATCTCGTTTCT-- 874

Qy 550 AAGTCCGACCGCAAGTCCAAAGTATGTGATTTGGATACATCTGCGGTTGCTAGCTCAGCT 609
Db 875 -----CCAAGGAAAGTACATTAATTTGGTTTCTTATGTACCTTGGAGCTTCTGCA 924

Qy 610 GGTATTCTCTGCTGCTTCTTTTAAACAGATTAAGCGTTTCGAAAGATTTCTAAAGAAATAC 669
```


Db 925 GTGTAAGCTCTTTGCTTCCCTCATGCACTGACCTTTGAGAAGGTTCTGAAGAAGGAA 984
QY 670 ACATTCAAGGCTATTTTATGACATGGCCACATATCCGCTATAGTAGTACTCTGTGTAGTT 729
Db 985 ACATTTCTCTGTTGTTTGGAAATGCAAACTACACATCAATTCGTTGCTCTGCTGCTCT 1044
QY 730 GTGGTAGGACTTTTGGAAAGTGTGGGTGAAAGAGCTGAGTACAGAAATGGAAGAGTTT 789
Db 1045 GTCATAGGCCCTATTTGCAAGTGGGAAATGGCGTAATTGTCATGGAGAAATGGAGGGTTT 1104
QY 790 CAACCTAGGAAAAAGCTCATACATTTTGTATAAATCGGTTCAACGATATCATGCGAGCT 849
Db 1105 CAGAAAGGATATGTTGCTATATGATCTTGGTTGGACTTCAATAGCCTGGCAGGTA 1164
QY 850 TGTTCATGGAAGTGTGGTTTGAATATCGAAGTTTCAATCGCTTTTCCAAATGTCATA 909
Db 1165 TGCTCTGTTGGTGTGTTGGCTTGAATCTCTAGTGTCTCTCTACTCTCAATGTTATA 1224
QY 910 AGCACTCTTTGTTTACGAGTGTGCTGTTCTTGGCTGTGTTGCTCTTCCGCTGATGAGATG 969
Db 1225 AGCACAGTTTCTTTAGCCGTAACCTCTATTTGCTGTCTATAGTTTTCATGATAAGATG 1284
QY 970 AGTGAATCAAGTTGGTTGCAATGTTTTCGCAATCTGTCGCTATGCTTCTTATGCTTAT 1029
Db 1285 AATGGGGTGAAGATAATTTCTATGCTTTTGGCTCTATGGGTTTTCGCTCTTATATTTAT 1344
QY 1030 CAGCATTATGTCATATGATAGAAAGCCAGAAAGACCAAGAGCTTCTCAGTCTAAAGAA 1089
Db 1345 CAGAATATCTGATGATTCAGAGACACATGACATGACAGCTCTACTAAGTCCCAAT 1404
QY 1090 GA 1091
Db 1405 GA 1406

RESULT 6

US-10-416-898-9
; Sequence 9, Application US/10416898
; Publication No. US20040172670A1
; GENERAL INFORMATION:
; APPLICANT: Yale University
; APPLICANT: Walker, Elsiebeth
; APPLICANT: Dellaporta, Stephen
; TITLE OF INVENTION: MAIZE YELLOW STRIPE1 AND RELATED GENES
; FILE REFERENCE: 44574-5106-US
; CURRENT APPLICATION NUMBER: US/10/416,898
; CURRENT FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: PCT/US01/43101
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 60/249,222
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 83698
; TYPE: DNA
; ORGANISM: Zea mays
US-10-416-898-9

Query Match

Best Local Similarity 54.4%; Pred. No. 9.1e-57; Length 83698;
Matches 547; Conservative 0; Mismatches 438; Indels 21; Gaps 2;
QY 70 TGGAGGTTAAGAGTGTCTCTATGTCACCTCTCTAGCTGGAGAGACAATAGCCACT 129
Db 67228 TGGTGAATCTGTTTTCATAGCAATCTCTTCTCATCTCTGCTCAAGCCATTTCTGTT 67287
QY 130 CTCTTAGGTAGACTTTTACGAAAAAGGGCGGTAAAGCACATGGCTCGAAACCTTTGGTT 189
Db 67288 CTCTTTGGTCGGTTTTTATACAAATGAAGGTGGAACAGATAATGGATCTCTACTCTGTT 67347
QY 190 CAGCTTGTAGGTTTCTTTAAACCTTCTCTGCTATATATCTTAAGGCTGAGCGCTCC 249

Db 67348 CAAACCTGGGGCTTTCCCAATCTTTATCTCCCTCTTTCTCTCTCTCTCTCTCTCTCT 67407
QY 250 AAGACTAAAAACCATACCAAAAAAACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 309
Db 67408 TCTTCT 67458
QY 310 ATGGACTTTGGCTTGTCT 369
Db 67459 CT 67518
QY 370 CT 429
Db 67519 CT 67578
QY 430 TCTTACT 489
Db 67579 TATTTATACATCAATCT 67638
QY 490 ACCATATCT 549
Db 67639 TCTATCT 67696
QY 550 AAGTCGGCAGCCAACT 609
Db 67697 -----CTAAGTGGAGTTTACTTGAATGGGTGTTCTCTCTCTCTCTCTCTCTCTCT 67746
QY 610 GGTATTTCT 669
Db 67747 ATCTATCT 67806
QY 670 ACATTCAAGGCTATTTTATGACATGSCCACAATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 729
Db 67807 ACTCTCTCTATGTTTCTCTGAGATGCAAAATCTATACGCTGTTGCTCTCTCTCTCTCTCT 67866
QY 730 GTGGTAGGACTTTTGGAAAGTGTGGGTGGAAGAAAGCTGAGTACAGAAATGGAAGAGTTT 789
Db 67867 GTTATCGGATTTCTCGCAAGCGGGAATGGATGTTTGTGAGTGTGGAGATGGAAGAGTTT 67926
QY 790 CAACTAGGAAAAAGCTCATACATTTTGTATATAAATCGGTTCAACGATATCATGGCAAGCT 849
Db 67927 CAGGAAGGTCAAGTCAATTTATGTTTGTGACTTTGGTTCGGGCGAGCGGTTCTGTCATTTG 67986
QY 850 TGTTCATGTTGAAGTGTGGTTTGAATATCGAAGTTTCAATCGCTTTTTCCTAATGTCATA 909
Db 67987 GGTTCGTAGGAGCGGTCGCTTATATTTCTGGTGTCTCTCTCTCTCTCTCTCTCTCTCT 68046
QY 910 AGCACTCTTTGTTTACCAAGTGTGCTGTTCTTGTCTGTTGTTCTCTCTCTCTCTCTCTCT 969
Db 68047 AGTACGCTCTCACTCAATTTGTTACGCTCTCTCGCGGCAATTTGCGGTGTTCCATGACAAGCTG 68106
QY 970 AGTGGAAATCAAGTGTGTTGCAATGTTTTCGCAATCTCTGGGCAATCTCTGTTTCTTATGTTAT 1029
Db 68107 ACTGAGTTAAGATGTTGCGGATGCCATCGCTCTTACAGGATTCACGTTTATATCTCTAC 68166
QY 1030 CAGCATTATGTCATGATAGAAAGCCAGAAAGACCAAGAGCTTCT 1075
Db 68167 CAGAACTATCTTGATGACTTGAAGTACAAAGAGCAGCAGAGAGCTC 68212

RESULT 7

US-10-424-599-75161
; Sequence 75161, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 75161

; LENGTH: 997

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_38884C.1

US-10-424-599-75161

Query Match 19.0%; Score 217; DB 16; Length 997;

Best Local Similarity 56.0%; Pred. No. 1.7e-50;

Matches 490; Conservative 0; Mismatches 355; Indels 30; Gaps 3;

```
QY 61 AAGTACTCTTGGAGTTAAGAGTGTCTCTATGTCACTCTCCTTCTAGCTGGAGACA 120
DB 125 AAATACAAATGGTGCTACGTGTCCCTCTACATAATCTTTCTTGTAGGCCAATCT 184
QY 121 ATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAGCGGTAAAGACATGCTCGAA 180
DB 185 GCAGCCACTCTTTGGGAGATTATCTATGACAAATGGTGTATAGCAATGATGGCA 244
QY 181 ACCTTGGTTCAGCTTGTAGGTTTCTTTAAACCTTCTCTGCTATTAATCTT---AAAG 237
DB 245 ACATTTGTTCAATCAGCTGGATTCCCTGACTTCTTCCACTCTTTTACTTCCCAAGA 304
QY 238 CTTGAGCGGTCCAAGACTAAACCAATACCAAAAACACTACTTCTTC-----285
DB 305 CAAACACATGCAAGTTTAATAACAAATCCAGCAATATGATTAATCTACAAACAAA 364
QY 286 -----TTCTTGACACTATCTTTAGTGATATGGACTTGGCTTGTGCTGGACAT 339
DB 365 CCAAAATCTCCACCCTTGTTTCTCTATCTAGCTTTTGGCCTAATCTCACAGGGGAC 424
QY 340 TGTATTTGTACTCAATTTGGCTACTTTTACCTTCTCTGTCTCAACTTCTTTGATCTCT 399
DB 425 AACTTGATGATTTCTATGGACTTTTATATCTCCTCTTTCACCTATCTCTACTATGT 484
QY 400 GCGTGGCAATTTGGCTTTTAAAGCGGTCTCTTCTTACTTCTTAACTCACAATAATACA 459
DB 485 GCCACCAATATAGTCTTCAAACGGGTGTTCTTCTTCTCTCAATTTCCAGAAATTCACA 544
QY 460 CCATTTATACTCAATTTCACTTGTCTCTTAAACCATATCTTCTACACTTCTTGTATCCAA 519
DB 545 GCATTCATTAATCAATCTGTAGTCTCTTACCATAATCAGCTTCTCTGCTTGCATCAAC 604
QY 520 CATGAACAGAAATCTCCTCTTCTACTTCAAGTCCGAGCCAGTCCAAATGATGTGATT 579
DB 605 TCCGATTCGAGCGAGGAC-----TCAACCGGTCTTTCAGAGAAAGCATGTAAAT 655
QY 580 GGATACATCTGCGGGTCGGTAGCTCAGCTGGTTATTTCTCTGGTGTCTTTTAAACAGAT 639
DB 656 GGATTTCTTGCAACCAATGGTGCAATCAGCTACATCTCTTTGTACCTCTCTCTTGTGCA 715
QY 640 TACGGTTTCGAAAGATTCTAAAGAAATPACACATTCAGGCTATTTTAGACATGGCCACA 699
DB 716 CTTTCTTTTCAAGATTATAGAGAGAAACCTTTCTGCTGTGTGGACATGCAATTT 775
QY 700 TATCCGTCTATGGTAGCTACTTGTGTAGTGTGGTAGGACTTTTGGAGGTGGTGGTGG 759
DB 776 TACCACATCTTCAATGCTACATGTCTGTGTAGTAGGGTTGTTTGCAGGTGGAGAGTGG 835
QY 760 AAAAGCTGAGTACAGAAATGGAGAGTTTCAACTAGGGAAGCTCATACATTTTGATA 819
DB 836 AAAAGTTTGAAATAGATGAAGGATATGATGAAGGAAGTGTGGTGTGTTGATTATC 895
QY 820 AACATCGGTTCAACGATATCATGGCAAGCTTGTGTTGATTGGAAGTGTGGTGTGTTGATTATC 879
DB 896 GTACTATAGATTGCTGTGACATGGTATATATGCTCACTAGGATGCTAGGGGGGACGTTG 955
QY 880 GAAGTTTCATCGCTTTTTCATGTTTCCAAATGCTCATAGCAC 914
DB 956 GAGGAGTCTGCATGTTTCTCGATGTGAGAGTAC 990
```

RESULT 8

US-10-437-963-43860/c

; Sequence 43860, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 43860

; LENGTH: 2381

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(2381)

; OTHER INFORMATION: unsure at all n locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_46977C.1

US-10-437-963-43860

Query Match

Best Local Similarity 17.2%; Score 197.2; DB 17; Length 2381;

Matches 528; Conservative 0; Mismatches 473; Indels 27; Gaps 2;

```
QY 70 TGGAGGTTAAGAGTGTCTCTATGTCACTCTCTCTAGCTGGAGAGACAATAGCCACT 129
DB 2246 TGGTGGCGATGGTGGCAGTGGACGCTCTTCTCTATCGCCGCTCAGACATCGGCGACA 2187
QY 130 CTCTTAGGTAGACTTTACTACGAAAAAGCGGTAAAAAGACATGGCTCGAAACCTTGTT 189
DB 2186 CTGCTGGCGAGATACTACTACCCAAAGCGCGCGAGCAAGTGGATATCGGCGTTCTGT 2127
QY 190 CAGCTTGTAGGTTTCTTTAAACCTTCTCTGTCTATTAATTAAGAGCTGAGCCCTCC 249
DB 2126 CGGACGCGCGGCTTCCCGATACTGTCTTCAACCTCTCTTCTTCCCTCGAAGTCACCG 2067
QY 250 AAGACTAAACCAATTACCAAAAAAACTACTTCTTCTTCTTGACACTATCTTTAGTGTAT 309
DB 2066 TCCTCTGCACCAACA-----CTCCCATGGCCAAGCTCGCGGTATATAC 2022
QY 310 ATTGGACTTGGCTTGTGTGTGGACATTTGATTTTGTACTCAITTTGGGCTACTTTAC 369
DB 2021 ATCGTCTCGGCTCATCATCGCGCGAGACATGATGATACACTGGTGGCTCAAGTAC 1962
QY 370 CTTCTGTCTCAACTTCTCTTTTGATCTCTCGCTCGCAATTTGCTTTTAAAGCGCTTTC 429
DB 1961 CTCCCGCTTCGACCTACTCGCTACTGTGTGCCAGCAGCTCGCGTTCAATGTCTTTC 1902
QY 430 TCTTACTTCTAAACTCACAATAAAATCACCATTATATCTCAATTCACATTTCTCTTA 489
DB 1901 TCATAGTGTCTCACTCCGAGAGGTCCTCTCTGATATTTCACTCGCTCGTCTGCTC 1842
QY 490 ACCATATCTTCTACACTTCTTGTATCCAAACATGAACCAAAATCTCCTCTTCTACTTCA 549
DB 1841 ACCATGTCGCTTCGCTCATCGGAGTCAGCAAGGA-----GTCTCAGGGGGTTC 1794
QY 550 AAGTCCGAGCGCAAGTCCAGTATGATTTGGATACATCTCGCGGTGCTAGTCTCAGCT 609
DB 1793 ACCGCGTCTCGGAGGGAAGTATCTGCTCGGTTTCTGCTGAGCGTGGGGGCGCTCGTGC 1734
QY 610 GGTATTCTCTGGTGTCTTTCTTTAAACAGATTACGCGTTTCGAAAGAGATTCTTAAAGAAATAC 669
```

Db 1733 ACCTACTCGTGTGATTCCTCGCGCTGATGACGCTCACCTTCGAGACCATCATCAAGAAGCAC 1674
Qy 670 ACATTCAAGAGCTATTTTAGACATGGCCACATATCCGCTCTATGCTAGTACTCTTGAGTT 729
Db 1673 ACCTTCTCAGCGCTCTCAACATGCAGATCTACACGCGCTCTGCGGAGCGCGGTG 1614
Qy 730 GTGGTAGGACTTTTGGAAAGTGTGGTGGAAAGCTGAGTACAGAAATGGAAGATTT 789
Db 1613 GTGGTCGGGCTGTTTCGCGAGCGGAGTGGAGTCTGAGGGGGGAGATGAACGCGTTC 1554
Qy 790 CAACTAGGAAAGCTCATACATTTTGTATTAACATCGGTTCAACGATATCATGGCAAGCT 849
Db 1553 AGTTCGGGGCAGTTCCTACCTGATGACGCTCTGCGGCGCGCTGCTGTCGGCAGGTG 1494
Qy 850 TGTTCGATGGAAGTGTGTTGTTGATTATCGAAGTTTCATCGCTTTTTCCTCAATGTCATA 909
Db 1493 GCCAACATCGGGGTGCTGGCCCTCATCTTCGAGTGTGCGGCTCTTCTTCCACGTCATC 1434
Qy 910 AGCACTCTTTGTTTACAGTTGTCCTGTTCTTGTGCTGTTGTTCTTCTTCCGTGATGATG 969
Db 1433 AGCACGGTGTCACTGCGCGTTCATCCCGTTCCTCGCGTGTGTTTCCACGACGATG 1374
Qy 970 AGTGGATCAAGTTGGTTCGAATGTTTTCGGCCATCTGGGATTTGTTCTTATGTTAT 1029
Db 1373 AACCGGGTGAAGATCGTGGCCATGCTGATTCGAATTTGGGATTTATTTTCGTATCTGTT 1314
Qy 1030 CAGCATTTGCTCAATGATAGAAAGCCAGAAAGACCAAGAGCTTCTCAGTCTAAGAA 1089
Db 1313 CAGCATTTGATGCGAAGAAAGCGAAGAGGCTTCATCGGGTGATGATGTCGAGG 1254
Qy 1090 GAAGAAGA 1097
Db 1253 CAAGAAGA 1246

RESULT 9

US-10-767-701-9000
; Sequence 9000, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 9000
; LENGTH: 728
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS69833_1
US-10-767-701-9000

Query Match 14.7%; Score 168.6; DB 17; Length 728;
Best Local Similarity 60.7%; Pred. No. 7.3e-37;
Matches 276; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

Qy 598 GGATGCTCAGCTGTTTCTCTGGTGTCTTTTAAACAGATTACGCGTTTCAAGAGATT 657
Db 1 GGAGCTCAGCCACATCTACTAATCTCTCTAATGCAAGTCAATTTGAGAGTT 60
Qy 658 CTAAGAATAACATTTCAAGGCTATTTTAGACATGGCCACATATCCGCTATGTTAGCT 717
Db 61 ATTAAGAAGAAAGCTTCTCAGTCGTTGTAATATGCATATATACACACTAGTGGA 120
Qy 718 ACTTGTGTAGTTGTGTAGACTTTTGGAAAGTGTGGTGGAAAGCTGATGACAGAA 777
Db 121 ACAATAGCTTCTCTTGTGGTTATTTGCAAGCGGTGAGTGGAAAGACTTTAGAGGAGAG 180

Qy 778 ATGGAGAGTTTTCAACTAGGAAAAGCTCATACATTTTGTATAAACATCGGTTCAACGATA 837
Db 181 ATGCATGCCCTTCAGCTCAGGAGGTTGTCTATGTGATGACACTTCTATGGAGCTGCTGA 240
Qy 838 TCATGGCAAGCTGTTTGTATGGAAGTGTGTTGATTTATCGAAGTTTCATCGCTTTT 897
Db 241 TCTTGGCAGATAGCTTCTGTAGGAGTGGTGGTTTGTATCTTTGTTGTGTCATCACTCTTT 300
Qy 898 TCCAATGTCAAGCACTCTTTGTTTACCAGTTGTGCTGTTCTTGTCTGTTCTTCTTC 957
Db 301 TCAATGTGATGAACACACTAGCTCTACCCATCATTTCCAGTTTTTGTCTGATTTTCTTC 360
Qy 958 CGTGATGAGATGGAATCAAGTTGGTTGCAATGTTTTCGCCCATCTGGGATTTGTT 1017
Db 361 CACGACAAGATGAATGAGTGAAGATTATGCTATGTTGATGCCATCTGGGATTTGTT 420
Qy 1018 TCTTATGGTTATCAGCATTTATGTCATGATAGAGAA 1052
Db 421 TCATATGGATACCAATTATATATTAGTGACAAGAA 455

RESULT 10

US-10-767-701-6806
; Sequence 6806, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 6806
; LENGTH: 747
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS55663_1
US-10-767-701-6806

Query Match 14.2%; Score 163; DB 17; Length 747;
Best Local Similarity 60.5%; Pred. No. 2.9e-35;
Matches 268; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

Qy 69 TTGGAGGTTAAGAGTGTCTCTATGTCACTCTCTCTTAGCTGAGAGACAATAGCCAC 128
Db 304 TTGGTGGTTGATGGTAGCTCTGAACATTTTCTCTCTGTTGCTGCTCAGACATCGAC 363
Qy 129 TCTCTAGGTAGACTTTTACTAGAAAAAGGCGGTAAAAAGCATGCGTGGAAACCTTGGT 188
Db 364 ACTCTCTGGTAGGTTCTACTACAAAGGTGGAACAGCAATGGATGTCCACATTTGT 423
Qy 189 TCAGCTTGTAGGTTTCTTTAACCTTCTTGTCTTATTTACTTAAAGCTGAGCCGTC 248
Db 424 CCAAAACCGCTGGCTTCCAGTATTTGTTGTCGCCCTATATCTGTTCCGTTCAAGTCGCC 483
Qy 249 CAAGACTAAAAACATTTACAAAAAACTACTTCTTCTTCTTGACATCTTTTAGTGTA 308
Db 484 TTCTACTCAAAACACACCAAGTAACCTGAGACTTCTGTCCACAGATTTACTTTATATA 543
Qy 309 TATTGGACTTGGCTTGTGTTGCTGGACATTTGTAATTTTGTACTCATTTGGGCTACTTTA 368
Db 544 CATTTGGCTTGGGCTCATCATTTGCTGCGATGACTTGATGTATTTCTATGGCTCTTGTGA 603
Qy 369 CTTTCTCTCAACTTTCTCTTTGATCTCTGCGTCGCAATTTGGCTTTTAAAGCCGCTCT 428
Db 604 CTTTCTCTGATCAACATATTCGCTCATATGCGCTAGTCAGCTGGGCTTCAATGCTGCTT 663
Qy 429 CTCTTACTTCTTAAACTCACAATAATCACACCAATTTATATCTCAATTCACTTGTTCTCTT 488

Db 664 CTCCTTATGTCCTGAATGCTCAAAAGTTCACTGCAATGATCTCAACTCGGTAAATCCTCT 723
Qy 489 AACCATATCTTACACTTCTTG 511
Db 724 TACCTTTTCTGTCGCACTTCTTG 746

RESULT 11
US-09-770-152-44
; Sequence 44, Application US/09770152
; Publication No. US20020040489A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: Thaliana
; FILE REFERENCE: 2025US (PARA-014PRV)
; CURRENT APPLICATION NUMBER: US/09/770,152
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,503
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-152-44

Query Match 13.5%; Score 154.6; DB 9; Length 592;
Best Local Similarity 56.9%; Pred. No. 6e-33;
Matches 283; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

Qy 28 GAAGGGAATTTTCAGAGAGAGAGAGTCAAGTACTCTTGAGGTTAAGAGTGTCT 87
Db 94 GAAAGTGAATCATTTTTCAGTACCTCAACGAGAACTGTAAAGAGTGGCTCCGTGCTCC 153
Qy 88 CTCATATGTCCTCTCTAGCTGGAGACAATAGCCACTCTCTTAGGTAGACTTTAC 147
Db 154 ATATACGCAATCTTGTCACTCTTGCACACCTTGTCTACAGTTCTGGGTAGACTGTAC 213
Qy 148 TACGAAAAGGGGTAAAGCACATGCTCGAAACCTTGGTTCAAGCTTGTAGGGTTTCCT 207
Db 214 TATGAAATGAGGGAAGAACACATATGTTGTAACACTTCTTCAACTCATTTGGCTTCCT 273
Qy 208 TTAACCTTCTCTGTATATATCTTAAAGCTGAGCGGTCCAGACTAAACATATACC 267
Db 274 GTACTGATCTCTTCGGCTCTTTTCTCGAATCAGGCAACCAATCAACAGATACAAAT 333
Qy 268 AAAAAAACTACTCTCTCTCTGACACTATCTTTAGTGTATATTTGAGCTTGGCTTGTCT 327
Db 334 TTCAGTCACTCCCTTCTCTTCAACACCTTCGATCGGTTTACTTGTGCACTGGACTGTA 393
Qy 328 GTTCTGGCAATGTATTTTGTACTCAFTTGGGCTACTTTTACCTTCCTGTCTCAACTTTC 387

Db 394 GTGTCTGCTTATGCTTATTTGTTCTGCAAGTTGGGTGCTTTACTTACCAGTCTCTACTTTC 453
Qy 388 TCTTTGATCTCTGCTCGCAATGGCTTTTAAAGCGCTCTCTCTTACTTACTTAACTCA 447
Db 454 TCCTCATCTTGGCCTCACAGTTGGCCTTCACTGCCCTTTTCTCATATTTCCCTTAACTCG 513
Qy 448 CAAAAAATCACACCATTTATATCAATCACTTCACTTCTTCTTAAACATATCTTCTACACTT 507
Db 514 CAAAAGTTCACTCTTTGATAGTCAATTTCTTTGTTCTCTCTAGGTTTCTCTGCGCTC 573

Qy 508 CTTGTTATCCAACATGA 524
Db 574 CTGCTGGTCAACACTGA 590

RESULT 12
US-10-437-963-83900
; Sequence 83900, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 83900
; LENGTH: 2605
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_83188C.1
US-10-437-963-83900

Query Match 11.4%; Score 130.6; DB 17; Length 2605;
Best Local Similarity 49.7%; Pred. No. 9.2e-26;
Matches 368; Conservative 0; Mismatches 364; Indels 9; Gaps 1;

Qy 305 TGTATATTGACTTGGCTTGGCTTGGCTGGACATTTTGTACTCAATTTGGGTAC 364
Db 1526 TCTACGTGCGCTCGGATCATCATCGGCTTCGACAACTGATGTACTCGTACGCGTGC 1585
Qy 365 TTTACCTTCTGCTCAACTTCTCTTTGATCTCTGCTCGCAATTCGGCTTTTAAACCGC 424
Db 1586 AGTACCTGCGCGGTGTCCACCTTCTCGCTCGTGGCGCGACGAGCTCGGCTTCAACTCCG 1645
Qy 425 TCTTCTTACTTCTTAACTCAAAAAAATCAACCAATTTATCTCAATTTCACTTGTTC 484
Db 1646 TCACCTCCAGCTCATCAACGCGAGCGGTTTCAAGGTGCTGATCGCAACTCCGTCGTCG 1705
Qy 485 TCTTAAACATATCTTTCATCTTCTTTGATATCCAAATGAACCAAGATCTCCCTTCTTA 544
Db 1706 TGCTCACCTTCTCGGCGCGCTGCTCGGCATC-----GGGGCCTCTCTCCGACGAGA 1756
Qy 545 CTTCAAGTCCGAGCCCAAGTCCAAAGTATGTGATTGATACATCTGCGCGTTCGGTAGCT 604
Db 1757 CCGCAGACGCTGCCGAGGGCAAGTACCCCGCGGGTTGCGCCTGACGCTTCGCCGCT 1816
Qy 605 CAGCTGTTTATCTCTGGTGTCTTTTAAACAGATTACGGGTTTCGAAAAGATTCTAAGA 664
Db 1817 CCGCGTGTTCGCGCTCATCTCTGCTGTTTCGAGGCCACTTCGAGAAGGTGTCGCGGA 1876
Qy 665 AATACACATTCAGGCTATTTTATAGACATGCGCACATATCCGTCTATGGTAGTACTTGTG 724
Db 1877 CGCGAGCTCCGCTGGGTTCTCGCGCGCAGCTGTGGACCAACGCTGTCGCGCTCGACGG 1936


```
Qy 707 CTATGCTAGCTACTTGTGTAGCTTGTGTAGGACTTTTGGAAAGTGTGGGTGGAAGAAGC 766
Db 1168 ACCTGGTGGCGTGGCGGTGTCCGTGGCGGGCTGCTCGCTCGGGGGACTTGGCGGACGA 1227
Qy 767 TGAGTACAGAAATGGAAGAGTTTCAACTPAGGGAAGCTCATACATTTTGATAAACAATCG 826
Db 1228 TCCCGGGGAGATGGCGTCGTTCAAGACGGGAGGACGAGGTACGTGCTGACGCTGGTGG 1287
Qy 827 GTTCAACGATATCATGGCAAGCTGTTTGAATGGAAGTGTGGTGTGATTAATCGAAGTTT 886
Db 1288 GCACGGCGGTGTGGTGGCAGCAGCGCGGTGGGCTTGGTGGCGCTGATCATGAGGGTGT 1347
Qy 887 CATCGCTTTTCCAAATGTCATAAGCACTTTTGTACCAAGTGTGCGCTGTTCTGCTG 946
Db 1348 CGTGGCTGTTCCGGAACGTGACGTGCACGCTGGCGCTGCCGCTGGTCCAGTGTTCGG 1407
Qy 947 TTGTCTTCTCCGTGATGAGTGAATCAAGTTGGTGTGCAATGTTTTTGGCCATCT 1006
Db 1408 TGGCGCTGTTGGGGACAGGATGACCGGTATAAGATCGTGGCCATGCTCATGGGGTAC 1467
Qy 1007 GGGGATTTGTTCTTATGTTATCAGCAATATGTCOAATG 1045
Db 1468 GGGGTTCTCTCTCTACATGTACCAAGCACTACATCGACG 1506
```

RESULT 15

```
US-10-425-114-12485
; Sequence 12485, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 12485
; LENGTH: 1386
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 701179255_FLI
US-10-425-114-12485
```

```
Query Match 11.0%; Score 125.4; DB 16; Length 1386;
Best Local Similarity 49.9%; Pred. No. 1.9e-24;
Matches 349; Conservative 0; Mismatches 341; Indels 9; Gaps 1;

Qy 347 TGTACTCATTTGGGCTACTTTTACCTTCTGCTCAACTTCTTGTGATCTCTGCGTCG 406
Db 453 TGTACTCATGCGCTGAGTACCTTCCCGGTGTCACCTTCTCGCTGCGCGGACGC 512
Qy 407 AATTGGCTTTTAAACGCCGCTTCTTCTTACTTCTTAAACTCAAAATAATCACCATTTA 466
Db 513 AGCTGGGCTTCAACGCCATCACCCTCGCGCTCATCAACGCGAGCGGTTCAOGGCTGA 572
Qy 467 TACTCAATTCAGTTGTTCTTTAACCATATCTTTACACTTCTTGTATCCAAATGAAC 526
Db 573 TCGCCAACTCCGTGGTGTGCTTCACTTCTCGCGCGATCTCGCGCTC-----G 623
Qy 527 CAGATCTCCCTTCTTACTTCAAGTCCGAGCCCAAGTCCCAAGTATGATGGATACA 586
Db 624 GATCTCTCTCGACGAGACCTTCCAGCAACGTGCGCGGCAAGTACCACCTGGGGTTCG 683
Qy 587 TCTGCGCGGTGCGGTAGCTCAGCTGTTATTTCTTGGTGTCTTCTTTAAACAGATTACGCGT 646
```

```
Db 684 TACTCGTGTGGCGCGCTCGCGGTGTTCGCGCTCATTTTGTCCCTCTTCGAGCTCACCT 743
Qy 647 TCGAAAGATTTTAAAGAAATACACATTTCAAGGCTATTTTAGACATGGCCACATATCCGT 706
Db 744 TCAGAAAGTGGTCAAGGTGCGACGGCGCGGTGGTGTGCGGATGCAGATGCACACCA 803
Qy 707 CTATGGTAGCTACTTGTGTAGTTGTGTAGGACTTTTGTGAAAGTGGTGGGTGGAAGAAGC 766
Db 804 ACCTGGTGGCGTGGCGGTGTCCGTGGCGGGCTGCTCGCTCGGGGACTGGCGGACGA 863
Qy 767 TGAGTACAGAAATGGAAGAGTTTCAACTAGGGAAGCTCATACATTTTGAATAAACAATCG 826
Db 864 TCCCGGGGAGATGGCGTGTTCAAAGACGCGGAGGACGAGGTACGTGCTGACGCTGGTGG 923
Qy 827 GTTCAACGATATCATGGCAAGCTTGTGTGATTTGGAAGTGTGGTGTGATTTATCGAAGTTT 886
Db 924 GCACGGCGGTGTCCGTGGCAGGACGCGCGCTGGGCTTGGTGGCGCTGATCATGAGGGTGT 983
Qy 887 CATCGCTTTTTCCAATGTCTAAGCACTCTTTTGTATTACCAAGTTGTGCCCTGTTCTTCTG 946
Db 984 CGTGGCTGTTTCGGAACGTGACGTGCACGCTGGCGCTGCCGCTGGTGGCAGTGTTCGG 1043
Qy 947 TTGTCTTCTTCCGTGATGAGATGAGTGAATCAAGTTGGTGTGCAATGTTTTTGGCCATCT 1006
Db 1044 TGGCGCTGTTTCGGGACAGGATGACCGGTATAAAGATCGTGGCAATGCTCATGGCGGTAT 1103
Qy 1007 GGGGATTTGTTCTTATGTTATCAGCAATATGTCOAATG 1045
Db 1104 GGGGTTTCTCTCTCTACATGTACCAAGCACTACATCGACG 1142
```

Search completed: November 1, 2004, 23:39:57

Job time : 599.039 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:48:03 ; Search time 105.63 Seconds
(without alignments)
7704.726 Million cell updates/sec

Title: US-09-913-767-3

Perfect score: 1145

Sequence: 1 cgtgttcattgtgtgtgtg.....caagtttaggaagatcca 1145

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/prodata/1/ina/5A COMB.seq.*

2: /cgn2_6/prodata/1/ina/5B COMB.seq.*

3: /cgn2_6/prodata/1/ina/6A COMB.seq.*

4: /cgn2_6/prodata/1/ina/6B COMB.seq.*

5: /cgn2_6/prodata/1/ina/PCUS COMB.seq.*

6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74.4	6.5	7218	1	US-08-232-463-14
2	45.6	4.0	399	4	Sequence 8976, Ap
3	45	3.9	3090	4	Sequence 5, Appl
4	42.6	3.7	3081	4	Sequence 9, Appl
5	41.6	3.6	1141	4	Sequence 22, Appl
6	41	3.6	1141	4	Sequence 22, Appl
7	40.2	3.5	500	3	Sequence 2, Appl
8	39.2	3.4	832	4	Sequence 2813, Ap
9	38.2	3.3	832	4	Sequence 2813, Ap
10	38.2	3.3	5935	3	Sequence 17, Appl
11	38.2	3.3	5935	3	Sequence 29, Appl
12	38.2	3.3	5935	3	Sequence 29, Appl
13	36.8	3.2	843	4	Sequence 5526, Ap
14	36.8	3.2	5092	4	Sequence 153, App
15	36.6	3.2	19233	4	Sequence 45, Appl
16	36.2	3.2	582	4	Sequence 3, Appl
17	35.8	3.1	1608	4	Sequence 1772, Ap
18	35.8	3.1	14759	4	Sequence 1, Appl
19	35.6	3.1	3117	4	Sequence 275, App
20	35.4	3.1	1447	4	Sequence 27, Appl
21	35.2	3.1	289	3	Sequence 17, Appl
22	35.2	3.1	289	3	Sequence 17, Appl
23	34.8	3.0	10619	4	Sequence 4, Appl
24	34.6	3.0	10007	3	Sequence 13, Appl
25	34.4	3.0	573	3	Sequence 440, App
26	34.4	3.0	789	4	Sequence 1652, Ap
27	34.4	3.0	879	4	Sequence 87, Appl

28	34.4	3.0	933	4	US-09-107-532A-78	Sequence 78, Appl
29	34.4	3.0	18073	3	US-09-078-294-12	Sequence 12, Appl
30	34.2	3.0	455	4	US-09-232-785-296	Sequence 296, App
31	34.2	3.0	722	4	US-09-270-767-8029	Sequence 8029, Ap
32	34.2	3.0	722	4	US-09-270-767-23311	Sequence 23311, A
33	34.2	3.0	1395	4	US-09-134-000C-2349	Sequence 2349, Ap
34	34.2	3.0	2834	1	US-08-276-151-8	Sequence 8, Appl
35	34.2	3.0	3054	4	US-09-762-724-11	Sequence 11, Appl
36	34.2	3.0	3073	1	US-07-688-352C-31	Sequence 31, Appl
37	34.2	3.0	3073	2	US-08-474-379C-31	Sequence 31, Appl
38	34.2	3.0	3073	3	US-09-146-249A-31	Sequence 31, Appl
39	34.2	3.0	3073	3	US-08-206-188B-31	Sequence 31, Appl
40	34.2	3.0	3073	5	PCT-US91-02714-30	Sequence 30, Appl
41	34.2	3.0	3084	4	US-09-762-724-7	Sequence 7, Appl
42	34.2	3.0	8093	4	US-10-204-708-32	Sequence 32, Appl
43	34	3.0	747	4	US-09-134-000C-2966	Sequence 2966, Ap
44	34	3.0	1170	4	US-09-248-796A-9079	Sequence 9079, Ap
45	34	3.0	1298	3	US-08-948-705-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHREIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F15
US-08-232-463-14

Query Match 6.5%; Score 74.4; DB 1; Length 7218;

TITLE OF INVENTION: Identification of a region of the major surface
; TITLE OF INVENTION: glycoprotein (MSG) gene of human Pneumocystis carinii
; FILE REFERENCE: 4239-58054
; CURRENT APPLICATION NUMBER: US/09/762,724
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: PCT/US99/18750
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: US 60/096,805
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3081
; TYPE: DNA
; ORGANISM: Pneumocystis carinii sp. f. hominis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3030)
US-09-762-724-9

Query Match 3.7%; Score 42.6; DB 4; Length 3081;
Best Local Similarity 46.2%; Pred. No. 0.031;
Matches 141; Conservative 0; Mismatches 164; Indels 0; Gaps 0;
QY 206 CTTTAAACCTTCCTTCGTTATTTACTTAAAGCTGAGCGGTCCTCAAGACATAAACCATTA 265
Db CCTTAACTCTGCAATCTGATTCCTCAATTTATTACATCTTTTCAAGCTCTAGATATCTTC 2624
QY 266 CCAAAAAAATCTCTCTCTCTGACACTATCTTTAGTGATATATGACATGCTGCTCAACTT 385
Db CAAATCTCTGCTCAAAATCAATCAATGCTTACCTTCTGCTCAACTT 385
QY 326 TTGCTGCTGACATTTGTTTCTGCTCAATTTGGGCTACTTTTACCTTCTGCTCAACTT 385
Db TAACTCTCTGCTGATTTTCTGCTCAATTTGGGCTACTTTTACCTTCTGCTCAACTT 385
QY 386 TCTCTTGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 445
Db TATCTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 444
QY 446 CACAAAAATCAACACATTTATCTCAATTCATCTGCTTCTCTTAAACCATATCTTCTACAC 505
Db ATGATAAACAAGATGGATTTATTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
QY 506 TTCTT 510
Db 2383 ATGTT 2379

RESULT 5
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEl promoters
US-09-806-708B-22

Query Match 3.6%; Score 41.6; DB 4; Length 1141;
Best Local Similarity 9.5%; Pred. No. 0.036;
Matches 64; Conservative 278; Mismatches 327; Indels 8; Gaps 2;
QY 472 AATTCACTTGTCTCTTAACCATATCTTCTACACTTCTTGTATATCAACATGACAGAA 531
Db 70 RWNKSRWKGWYKKWYBCANNYSBRYHARRWKDMKTAYBMTMTNKGKGTGWRHYWR 129
QY 532 TCTCCTCTTCTACTTCAAAAGTCCGAGCAGCAAGTCAAGTATGTGTGATGATCTGTC 591
Db 130 WRAMBDTVDHVYVAMNNATWTCMDKDDKTRWKKNNNATGDDDTKHHMNNNGC 189
QY 592 GCGTCGGTAGCTCAGCTGCTGTTATTTCTCTGCTCTTTTAAACAGATTACCGGTTGAA 651
Db 190 BTVTWVRYKTDWDWSBKRMNYGBWKNWSYDVTVYVWVWDDMKCRKRVVRVTRGRMR 249
QY 652 AAGATTTCTAAAGAAATACACATTTCAAGGCTATTTTAGACATGCGCCACATATCCG 705
Db 250 NYWVAMBTARRRRYNNGMTBAMAYRRWNNNNNAKAMCKRAKYGMWRABVNSTCTTW 309
QY 706 TCTATGCTAGTACTTGTGTAGTGTGTGAGGACTTTTGGGAAGTGGTGGTGAAGAAAG 765
Db 310 KSKTKYRTSCHWNNCRAGDANKDHKWKWSAAMGVYNNNNNNNNNNNNNNNNNNNNNN 369
QY 766 CTGAGTACAGAAATGAAGAGTTTCAACTAGGAAAGCTCATACATTTTGTATAAATC 825
Db 370 MHSWKWKHANAHAHYSRKKWTBYKRTVMNNNGTTTMMKRWAWYWKMDWBGTYNNNN 429
QY 826 GGTTCACAGATATCATGCAAGCTTGTGTTGATTTGGAAGTGTGTTGATTTGATTCGA 885
Db 430 NGRTYTYGWTGKWKWMTYKWKANNCKWRADHKTCTHNTTTWKKMTYNNNNNNNNNN 489
QY 886 TCATCGCTTTTTCCTCAATGCTAAGCACTCTTGTGTTTACCAAGTGTGCTGCTGCTGCT 945
Db 490 GSHSRBAAYVYTWVWRRYAHANNNDWYKWKACTYKYBVCKWNNNNYAAWYTKSSW 549
QY 946 GTTGTCTTCTTCGCTGATGAGTGAATCAAGTGTGTTGCTCAATGTTTGGCCATC 1005
Db 550 NTSRYRWKTNWSRWSRSDTRSMGRANNYARABHYGKYNTRWBSHTWBHBRAGAAH 609
QY 1006 TGGGATTTGTTTCTTATGTTATGATGATGATGATGATGATGATGATGATGATGATG 1063
Db 610 YMBWVYBAKCHMKAWYKAKYAGAGGSSNNNNNNNNNNNNNNNNNNNNNNNNNNNN 669
QY 1064 ACCAAGAGCTTCTCAGTCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1123
Db 670 AMANAKWYKBAANNAYTHANNWGWGNNATDTRTWKNNNNNNNNNNNNNNNNNN 729
QY 1124 TCCAAGCTTAGGCAAG 1140
Db 730 NASAANKYAAAAVKA 746

RESULT 6
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter

```
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEI promoters
US-09-806-708B-22

Query Match          3.6%; Score 41; DB 4; Length 1141;
Best Local Similarity 10.5%; Pred. No. 0.054;
Matches 105; Conservative 344; Mismatches 533; Indels 14; Gaps 1;

QY 2 TGTTCATGTTGATAGTAGAAGCTGAAGGAAATTTTCAACAGAGAGAGAGAGTGCACA 61
DB 1126 TKTYKANNNNNNNGMGKDNRMDDATKWSATGTAWTTHAKRGATMCWYWTGTNRW 1067
QY 62 AGTACTCTGAGGTTAGAGTGCTCTCTATGACACTCTCTCTAGCTGAGAGACAA 121
DB 1066 CMRTYAMRTWYTRSNANWSCATKBNWMTWKYATKYRTAWYAMCAMNNNNWCATNGYA 1007
QY 122 TAGCCACTCTCTAGTAGACTTTACTACGAAAAAGCGGCTAAAAGACATGGCTCGAAA 181
DB 1006 KSCATNNAMWATTFWAAYAAKWARWAGNNRMYGAAAGNKWGCMAAMATMBGWADTA 947
QY 182 CTTGGTTCAGCTGTAGGGTTCTTTAAACCTTCCTTGCTATATTACTTTAAAGCGTG 241
DB 946 GKMCNNNNNNWTTDVRMAKAKNNNNNNYWTACYNRAATNNKMWATHMWKWTHGAHSKR 887
QY 242 AGCCGTCCAGACTAAAACCATACCAAAAAAACTACTCTCTCTCTTGACACTATCTT 301
DB 886 RTRHRTTCRRTKYNNNNNNNARTVYVYHHAARRMWNAMWTRNNNNNNNNNACRNTW 827
QY 302 TAGTGTATATGGACTTGGCTTGTCTGTGGACATTTGATTGTGACTCATTTGGGC 361
DB 826 WABWKHSWNNNNNNNNNNNNNNWCHYTANABBCYBANNNAARMARTCNMYMHAUV 767
QY 362 TACTTTACTCTCTGCTCAACTTCTCTTGTGATCTCTGCGTCAATGGCTTTTAACG 421
DB 766 TTTHTDWCYKTMNTWYWDWMTTMBTTTTTRNNTSTNNNNNNNNNNNNNNNNNNMKA 707
QY 422 CGCTCTCTTACTCTTAACACTCACAACCAACACACCACTTATATCTCAATTCACCTG 481
DB 706 YYAHATNNWGCNNNTDARRNNTTVMRRNMTNKTWYSTTRHHYTGATNNNNNNNN 647
QY 482 TTCTCTTAAC-----ATATCTCTACACTCTTGTGTTATCCCAACTGAACC 527
DB 646 NNNNNNNSCCTCTEMTMRWTKMGDMTVRKVKWRDTCTTYDVWDADSWVWYANWR 587
QY 528 AGAATCTCCCTCTTACTCAAAAGTCGCGAGCCAGGTCGAAGTATGATGATGATACAT 587
DB 586 CRDVTYTRNTYCKSYAHSYWYWSNNAMWYRRYSARNWSSMARWTRNNNNWMSGBVRW 527
QY 588 CTGCGCGGTGAGTCACTGCTGTTATCTCTGCTGCTTCTTTAAACAGATTACGCGTT 647
DB 526 AGTMMRHHNNNTDTRYWYWKWARBTTVYDSMCAKSWRNNNNWRMKNMMAANNND 467
QY 648 CGAAAGAGTTCTAAAGAAATACACATCAAGGCTATTTTAGACATGGCCACATATCCGTC 707
DB 466 AGAMDHTWYMGNTTMMRRAKWMMAWCRRAVCNNNNNNRACVWHKHKWRWTKWYMW 407
QY 708 TATGTAGTACTCTGTGTAGTGTGCTAGGACTTTTTCGAAAGTGGTGGGAGAAAAGCT 767
DB 406 KAAACNNNNBKAMYRVAWMTYSRDTTNTDMMWTSDBWBHWYTVDTYVMRAWNNNNNNW 347
QY 768 GAGTACAGAAATGGAAGATTTCACATAGGGAAGCTCATACTTTTGTATGATAAATCGG 827
DB 346 RBCKTTSWMMWMDHNTCTGNTGNWGSAYBNMAAASMAAGASBNVTYVWCRWMTYMGKT 287
QY 828 TTCAACGATATCATCGCAAGTGTGTTGATTGGAAGTGTGTTGTTGATTGCGAAGTTTC 887
DB 286 MTNNNNNNKAWYRKTVAWCCNNRYDYTAVMTBKNKYCYAVBWBWYMYMGKHWW 227
QY 888 ATCGCTTTTTCATAGTACATAGCACTCTTTGTTTACAGTGTGCGCTGTCTGCTGTG 947
DB 226 RRABHRSNNMMWVKCRNKYMSWYHYHAMRYBKWABAVGNNNNWKKDRMAHHWCATNNNM 167
QY 948 TGCTCTCTTCGTGATGAGATGAGTGAAGTT 983
```

```
DB 166 WWWAYVHHMHKKGAANTNNKTABRDDHBAHVKT 131

RESULT 7
US-09-141-000-2/c
; Sequence 2, Application US/09141000
; Patent No. 6054295
; GENERAL INFORMATION:
; APPLICANT: Chen, Fang
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
; FILE OF INVENTION: RECEPTOR PROTEINS
; FILE REFERENCE: 19999V
; CURRENT APPLICATION NUMBER: US/09/141,000
; CURRENT FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human
; US-09-141-000-2

Query Match          3.5%; Score 40.2; DB 3; Length 500;
Best Local Similarity 8.9%; Pred. No. 0.06;
Matches 27; Conservative 103; Mismatches 175; Indels 0; Gaps 0;

QY 326 TTGTGTGTCGACATGTTATTTGTACTCATTTGGCTTAAAGCGGCTCTCTTACTTCTTAAACT 385
DB 388 YW...DY.S.RH..T.D..H..M.BT..H..RKSHSNT.T.TM.AB...M.BM.MMRY 329
QY 386 TCTTTGTATCTCTGCGTCGCAATTTGGCTTTTAAAGCGGCTCTCTTACTTCTTAAACT 445
DB 328 B....TYR....CT.YSD..HK.RH.TRB.MHHR..SYRB..C.....KWS...SK.HC. 269
QY 446 CACAAAAAATCACACCACTTATATCTCAATTCACCTTGTCTTAAACCATATCTTCTACAC 505
DB 268 S.SS.C..DMTWC..BB..YHT.HG.AA.TM.HC..KC...KTR.MH..TB..RSB.MA. 209
QY 506 TTCTTTGTTATCAACATGAACCAAGAAATCTCCCTCTTACTTTCAAAAGTCCGACCAAGT 565
DB 208 .MT..S...S.R.SS.SH.YYMRM.YCCYBYH.YBC.M.KCBM.GMK.YGT.GSMYYM 149
QY 566 CCAAGTATGTCATGGATACATCTCGCGGTGCGTAGCTAGCTGCTGTTATCTCTGCTGTC 625
DB 148 A..G.NAT.GSR..N..AYM..TWGT.GSTBCDRGST.HCGB.G.YM..TN..KR.GM. 89
QY 626 TTTCT 630
DB 88 TSH.K 84

RESULT 8
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
```



```
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 656962el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: PE_FL_Genes Version 1.0
; SEQ ID NO 153
; LENGTH: 5092
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (455)..(4195)
US-09-620-312D-153

Query Match 3.2%; Score 36.8; DB 4; Length 5092;
Best Local Similarity 48.8%; Pred. No. 2;
Matches 157; Conservative 0; Mismatches 162; Indels 3; Gaps 2;

QY 782 AAGAGTTTCAACTAGGAAAGCTCATACATTTTGATAAACATCGGTTCACCGATATCAT 841
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4523 AAGGTTTACATGCTTAAACACAGCTACCTTTGTAATACTTCATCTGATCAGAAGTG 4582
QY 842 GCGAAGCTTGTTGATGGAAGTGGTGGTTGATGATTCGAGTTTCATCGCTTTTTCCTCA 901
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4583 GTCATGCTGTTTGA--GATGGAGTTCGCTGCATTTTAGGACTATTGATACCTTTTAA 4640
QY 902 ATGCTAAGCACTCTTTGTTTACCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4641 TTGTTTATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4700
QY 962 ATGAG-ATGAGTGAATCAAGTTGGTGGCAATGTTTTGGCCATCTGGGATTTGTTCT 1020
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4701 TGAAGTTGAATGAAATAGGCTTAACCATCATGTTGACTCTCTCTCTCTCTCTCTCT 4760
QY 1021 TATGTTATCAGCATTTATGTAATGATAGAAAGCCAGAGAACCAAGAGTTCCTCAG 1080
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4761 GTTGGAAATTTGTTTATTTCTTTCATGTAATGCTTAATTTGAACAGATCACTTACTAAA 4820
QY 1081 TCTAAAGAAAGAAAGCAAAA 1102
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4821 GCITTAGAAGATGATTCANA 4842

RESULT 15
US-10-204-708-45
; Sequence 45, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 45
; LENGTH: 19233
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (76, 178..179, 273, 586, 648, 651, 920, 1014, 1173, 1197, 1228)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1264, 1267, 1559, 1607, 1677, 1703, 1742, 1848, 1850, 2033)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (2168, 2254, 2257, 2280, 2291, 2723, 3018, 3080, 4205, 4209)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (4216, 4225, 4239, 4246, 4254, 4265, 4270, 4274, 4288, 4294..4295)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (4298, 4363, 4610, 4627, 4675, 4693, 4698, 4738, 4777, 4781)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (4787, 4825, 4829, 4838, 4855, 5462, 5494, 5496, 5500, 5674)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (5690, 5697, 5705, 5715, 5890, 5910, 5913, 6027, 6080, 6109)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (6244, 6288, 6776, 6873, 6891, 6980, 7028, 7656, 7876, 8030)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (8059, 8075, 8632, 8636, 9700, 9717, 9815, 9818, 9867, 9879)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (10005, 10075, 10150, 10159, 10165, 10356, 11123, 11166, 11360)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (11703, 12050, 12340, 12981, 13564, 13764..13765, 13787)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (13846, 14168, 16566, 17006, 17217, 17650, 17653, 17656, 17898)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (18130, 18143, 18151, 18188, 18498, 18549, 18936, 18946, 19165)
; OTHER INFORMATION: n is a or g or c or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (19200)
; OTHER INFORMATION: n is a or g or c or t
US-10-204-708-45

Query Match 3.2%; Score 36.6; DB 4; Length 19233;
Best Local Similarity 50.3%; Pred. No. 4.7;
Matches 90; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 849 TTGTTTGATGGAAGTGTGTTTGCATTCATTCGAATTCGAATTCATCGCTTTTTCATGTCAT 908
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3477 TTTTAAGTTGGAGTGTAGTGGTTTGCATTTGTTTATGTAATTTTGTTCGGGTTT 3536
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:40:28 ; Search time 601.255 Seconds
(without alignments)
9996.736 Million cell updates/sec

Title: US-09-913-767-3

Perfect score: 1145

Sequence: 1 ctgtttcatgtgtgatggt.....caagtttaggcaagatcca 1145

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1145	100.0	1145	3	Aaa97921 A. thalia
2	1124.2	98.2	1292	3	Aac51119 Arabidops
3	1122.6	98.0	1295	3	Aac37289 Arabidops
4	318.4	27.8	3387	6	Abz14407 Arabidops
5	318.4	27.8	3387	8	Ada68246 Arabidops
6	317.2	27.7	1194	3	Aaa97923 A. thalia
7	308.6	27.0	1071	3	Aaa97925 A. thalia
8	304.8	26.6	2860	3	Aac47974 Arabidops
9	298.4	26.1	1047	6	Adg88190 A. thalia
10	298.4	26.1	1081	3	Aaa97924 A. thalia
11	294.4	25.7	1293	3	Aaa97922 A. thalia
12	243.2	21.2	83698	6	Abn85767 Arabidops
13	235.8	20.6	1403	3	Aac37652 Arabidops
14	190.2	16.6	1152	8	Ada70213 Rice gene
15	162.2	14.2	2175	8	Ada70748 Rice gene
16	154.6	13.5	592	10	Abx56692 Arabidops
17	130.6	11.4	1155	8	Ada69686 Rice gene
18	125.4	11.0	1040	3	Ada69688 Rice gene
19	112	9.8	1049	3	Aaa97920 A. thalia
20	104.4	9.1	332	12	Adp92185 Cotton ex
21	95.8	8.4	1225	3	Aaa97919 A. thalia

22	94.8	8.3	1173	12	ADN72520	Adn72520 Thale cre
23	92.8	8.1	2586	3	AAC44184	Aac44184 Arabidops
24	77.6	6.8	477	3	AAC37273	Aac37273 Arabidops
25	76.2	6.7	774	6	ABN99127	Abn99127 Arabidops
26	68.6	6.0	492	3	AAC36845	Aac36845 Arabidops
27	52.8	4.6	386	6	ABQ85215	Abq85215 Arabidops
28	50.4	4.4	493	3	AAC36831	Aac36831 Arabidops
29	49	4.3	13573	6	ABL33869	Ab133869 Human inm
30	46.4	4.1	446	3	AAC37188	Aac37188 Arabidops
31	45	3.9	3090	3	Aaz94065	Aaz94065 Pneumocys
32	44.6	3.9	17213	6	ABL33483	Ab133483 Human inm
33	44	3.8	15373	6	ABU32467	Abu32467 Human inm
34	43	3.8	6145	4	AAS46547	Aas46547 Tumour su
35	42.6	3.7	3081	3	Aaz94067	Aaz94067 Pneumocys
36	42.4	3.7	2322	5	ADM19363	Adm19363 Novel hum
37	42	3.7	402	6	ABZ08435	Abz08435 Human leu
38	41.8	3.7	94400	12	ADP08387	Adp08387 Human gly
39	41.6	3.6	476	5	ADL37109	Adl37109 Human ova
40	41.6	3.6	476	5	ADI71958	Adi71958 Human ova
41	41.6	3.6	2000	8	ADA71938	Ada71938 Rice gene
42	41.6	3.6	2185	6	AAL45822	Aal45822 Human sec
43	41.4	3.6	1418	3	AAC46395	Aac46395 Arabidops
44	41.2	3.6	3628	8	ABZ10117	Abz10117 Haematopo
45	40.6	3.5	555	4	ABA59777	Abas59777 Human foe

ALIGNMENTS

RESULT 1

AAA97921
ID AAA97921 standard; DNA; 1145 BP.

AC AAA97921;

DT 19-JAN-2001 (first entry)

DE A. thaliana PUP1 DNA #3.

KW PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.

OS Arabidopsis thaliana.

PN DE19907209-A1.

PD 24-AUG-2000.

PF 19-FEB-1999; 99DE-01007209.

PR 19-FEB-1999; 99DE-01007209.

XX (FROM/) FROMMER W.

PI Gillissen B, Buerkle L, Andre B, Frommer WB;

XX WPI; 2000-566202/53.

PT Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transport, encodes a nucleobase transporter protein of
Arabidopsis thaliana.

XX Claim 1f; Page 12-13; 24pp; German.

CC This invention describes a novel nucleic acid encoding a plant nucleobase
transporter (I). (I) is produced by complementation of a nucleobase
transport (NBT)-defective host cell with a plant gene bank by selection
of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
CC expression of (II) (when in antisense orientation), and to produce
transgenic crop plants. The transgenic plants have modified nucleobase
transport properties, e.g. altered affinity and substrate specificity

CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX
SQ Sequence 1145 BP; 309 A; 234 C; 226 G; 376 T; 0 U; 0 Other;

Db	901	AATGTCATAAGCACTCTTTGGTTTACCAAGTTGTGCCTGTTCTTGCTGTTGTCTTCCGT	960
Qy	961	GATGAGATGATGGAATCAAGTTGGTTGCAATGTTTTTGGCCATCTGGGATTTGTTTCT	1020
Db	961	GATGAGATGATGGAATCAAGTTGGTTGCAATGTTTTTGGCCATCTGGGATTTGTTTCT	1020
Qy	1021	TATGGTTATCAGCATTTATGTCAATGATAGAAAAGCCAGAAAGACCAAGAGCTTCCTCAG	1080
Db	1021	TATGGTTATCAGCATTTATGTCAATGATAGAAAAGCCAGAAAGACCAAGAGCTTCCTCAG	1080
Qy	1081	TCTAAAGAAGAAGAAAGAACAAAAACAAGTAGATACCAATTCATGTCCAAAGCTTAGGCAAG	1140
Db	1081	TCTAAAGAAGAAGAAAGAACAAAAACAAGTAGATACCAATTCATGTCCAAAGCTTAGGCAAG	1140
Qy	1141	ATCCA 1145	
Db	1141	ATCCA 1145	
RESULT 2			
AAC51119			
ID	AAC51119	standard; DNA; 1292 BP.	
XX	AC		
XX	AAC51119;		
DT	18-OCT-2000	(first entry)	
XX	XX		
DE	Arabidopsis thaliana	DNA fragment SEQ ID NO: 67344.	
XX	XX		
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX	XX		
OS	Arabidopsis thaliana.		
XX	XX		
FN	EP1033405-A2.		
XX	XX		
PD	06-SEP-2000.		
XX	XX		
DF	25-FEB-2000; 2000EP-00301439.		
XX	XX		
PR	25-FEB-1999;	99US-0121825P.	
PR	05-MAR-1999;	99US-0123180P.	
PR	09-MAR-1999;	99US-01233548P.	
PR	23-MAR-1999;	99US-0125788P.	
PR	25-MAR-1999;	99US-0126264P.	
PR	29-MAR-1999;	99US-0126785P.	
PR	01-APR-1999;	99US-0127462P.	
PR	06-APR-1999;	99US-0128234P.	
PR	08-APR-1999;	99US-0128714P.	
PR	16-APR-1999;	99US-0129845P.	
PR	19-APR-1999;	99US-0130077P.	
PR	21-APR-1999;	99US-0130449P.	
PR	23-APR-1999;	99US-0130510P.	
PR	28-APR-1999;	99US-0130891P.	
PR	28-APR-1999;	99US-0131449P.	
PR	30-APR-1999;	99US-0132048P.	
PR	30-APR-1999;	99US-0132407P.	
PR	04-MAY-1999;	99US-0132484P.	
PR	05-MAY-1999;	99US-0132485P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	06-MAY-1999;	99US-0132487P.	
PR	07-MAY-1999;	99US-0132861P.	
PR	11-MAY-1999;	99US-0134256P.	
PR	14-MAY-1999;	99US-0134218P.	
PR	14-MAY-1999;	99US-0134219P.	
PR	14-MAY-1999;	99US-0134221P.	
PR	14-MAY-1999;	99US-0134370P.	
PR	18-MAY-1999;	99US-0134768P.	
PR	19-MAY-1999;	99US-0134941P.	
PR	20-MAY-1999;	99US-0135124P.	
PR	21-MAY-1999;	99US-0135353P.	
PR	24-MAY-1999;	99US-0135629P.	

PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143442P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149360P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149428P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154035P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156598P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.

PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match

98.0%; Score 1122.6; DB 3; Length 1295;

Best Local Similarity 98.8%; Pred. No. 4.2e-295; Matches 1131; Conservative 0; Mismatches 14; Indels 0; Gaps 0;				
Qy	1	CTGTTCATGTGTTGATGGTAGAACCTCGAAGGGAAATTTTCAACAGAAAGAGAGAAGTCAC	60	
Db	53	CTCCATCTTCACGTAATGGCGAACCTCGAAGGGAAATTTTCAACAGAAAGAGAGAGNATCAC	112	
Qy	61	AAGTACTCTTTGGAGGTTAAGAGTGTCTCTCTATCTCACTCTCTCTCTTTAGCTGGAGAGACA	120	
Db	113	AAGTACTCTTTGGAGGTTAAGAGTGTCTCTATGTCTACTCTCTCTTTAGCTGGAGAGACA	172	
Qy	121	ATAGCCACTCTCTTAGGTAGACTTTACTAGAAAGAGGGGTAAAGCAATGGCTCGAA	180	
Db	173	ATAGCCACTCTCTTAGGTAGACTTTACTAGAAAGAGGGGTAAAGCAATGGCTCGAA	232	
Qy	181	ACCTTGGTTCAGCTGTGAGGGTTCCTTTAAACCCCTCTCTGTATTTACTTTAAAGCCT	240	
Db	233	ACCTTGGTTCAGCTGTGAGGGTTCCTTTAAACCCCTCTCTGTATTTACTTTAAAGCCT	292	
Qy	241	GAGCGGTCCAAGACTAAAAACATTTACAAAAAAACTACTTCTCTCTTTGACACTATCT	300	
Db	293	GAGCGGTCCAAGACTAAAAACATTTACAAAAAAACTACTTCTCTCTTTGACACTATCT	352	
Qy	301	TTAGTGATATTTGACCTTGGCTTGGTCTGTGTCGACATTTATTTTGTACTCAATTTGGG	360	
Db	353	TTAGTGATATTTGACCTTGGCTTGGTCTGTGTCGACATTTATTTTGTACTCAATTTGGG	412	
Qy	361	CTACTTTACCTTCCTGTCTCAACTTCTCTTTGATCTCTGGTCGCAATTTGGCTTTTAAAC	420	
Db	413	CTACTTTACCTTCCTGTCTCAACTTCTCTTTGATCTCTGGTCGCAATTTGGCTTTTAAAC	472	
Qy	421	GCGCTCTTCTTACTTCTTAAACTCACAAAAATCACACATTTATACTCAATTCACATT	480	
Db	473	GCGCTCTTCTTACTTCTTAAACTCACAAAAATCACACATTTATACTCAATTCACATT	532	
Qy	481	GTCTCTTTAACCATATCTTCTACACTTCTTGTATTACCAATGAAACAGAAATCTCCCTCT	540	
Db	533	GTCTCTTTAACCATATCTTCTACACTTCTTGTATTACCAATGAAACAGAAATCTCCCTCT	592	
Qy	541	TCTACTTCAAAGTCGCGAGCCAGTCCAAGTATGTGATTTGGATACATCTCGCGGTGGT	600	
Db	593	TCTACTTCAAAGTCGCGAGCCAGTCCAAGTATGTGATTTGGATACATCTCGCGGTGGT	652	
Qy	601	AGCTCAGCTGGTATTCTCTGGTCTTTCTTTAACAGATTACGGGTCGAAAAAGATTCTA	660	
Db	653	AGCTCAGCTGGTATTCTCTGGTCTTTCTTTAACAGATTACGGGTCGAAAAAGATTCTA	712	
Qy	661	AAGAAATACACTTCAGGCTATTTTAGACATGGCCACATATCCGTCTATGGTAGCTACT	720	
Db	713	AAGAAATACACTTCAGGCTATTTTAGACATGGCCACATATCCGTCTATGGTAGCTACT	772	
Qy	721	TGTGTAGTTGTGTAGACATTTTGGAAAGTGGTGGGTGGAAAAAGCTGAGTACAGAAATG	780	
Db	773	TGTGTAGTTGTGTAGACATTTTGGAAAGTGGTGGGTGGAAAAAGCTGAGTACAGAAATG	832	
Qy	781	GAAGAGTTTCAACTAGGAAAGCTCATACATTTTGTATAACATCGGTCGAAAGATATCA	840	
Db	833	GAAGAGTTTCAACTAGGAAAGCTCATACATTTTGTATAACATCGGTCGAAAGATATCA	892	
Qy	841	TGGCAAGCTTGTTTGTATTGGAAGTGTGGTTTGTATTATCGAAGTTTCATCGCTTTTTTCC	900	
Db	893	TGGCAAGCTTGTTTGTATTGGAAGTGTGGTTTGTATTATCGAAGTTTCATCGCTTTTTTCC	952	
Qy	901	AATGTCTATAGCACTCTTTGTTTACAGTTGTGCTGTTCTGCTGTGTCTCTTCCTCCGT	960	
Db	953	AATGTCTATAGCACTCTTTGTTTACCAAGTTGTGCTGTTCTGCTGTGTCTCTTCCTCCGT	1012	
Qy	961	GATGAGATGAGTGAATCAAGTTGGTTCGCAATTTTTTGGCCATCTCGGGATTTGTTTCT	1020	
Db	1013	GATGAGATGAGTGAATCAAGTTGGTTCGCAATTTTTTGGCCATCTCGGGATTTGTTTCT	1072	
Qy	1021	TATGGTTTATCAGCAATTATGTCAATGATAGAAAGCCAGAAAGACCAAGAGCTTCCTCAG	1080	

Db	1073	TATGGTTATCAGCATTTATGTCAATATATAGAAAAGCCAGAGAAAGACCAAGAGCTTCCTCTCAG	1133
Qy	1081	TCATAAAGAAAGAAAGAAACAAAAACAAGTAGATACCACTTCATGCTCCAAGCTTTAGGCAAG	1140
Db	1133	TCATAAAGAAAGAAAGAAACAAAAACAAGTAGATACCACTTCATGCTCCAAGCTTTAGGCAAG	1192
Qy	1141	ATCCA 1145	
Db	1193	ATCCA 1197	
RESULT 4			
ABZ14407	ID	ABZ14407 standard; DNA; 3387 BP.	
XX	AC	ABZ14407;	
XX	DT	21-JAN-2003 (first entry)	
XX	DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 2212.	
XX	KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.	
XX	OS	Arabidopsis thaliana.	
XX	PN	WO200216655-A2.	
XX	PD	28-FEB-2002.	
XX	PF	24-AUG-2001; 2001WO-US026685.	
XX	PR	24-AUG-2000; 2000US-0227866P.	
XX	PR	26-JAN-2001; 2001US-0264647P.	
XX	PR	22-JUN-2001; 2001US-0300111P.	
XX	PA	(SCRI) SCRIPPS RES INST.	
XX	PA	(SYGN) SYNGENTA PARTICIPATIONS AG.	
XX	PI	Harper JF, Kreps J, Wang X, Zhu T;	
XX	PI	WPI; 2002-304127/34.	
XX	PT	Identifying a stress condition to which a plant cell has been exposed and	
XX	PT	producing plants with increased tolerance to these abiotic stresses.	
XX	PS	Claim 144; SEQ ID NO 2212; 577pp + Sequence Listing; English.	
XX	CC	The invention relates to identifying a stress condition to which a plant	
XX	CC	cell has been exposed, comprising: (a) contacting nucleic acid	
XX	CC	representative of expressed polynucleotides in the plant cell with an	
XX	CC	array or probes representative of the plant cell genome; and (b)	
XX	CC	detecting a profile of expressed polynucleotides in the plant cell	
XX	CC	characteristic of a stress response. The method is useful in the	
XX	CC	production of transgenic plants, cells and seeds and in producing plants	
XX	CC	with increased tolerance to abiotic stress. The present sequence is that	
XX	CC	of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used	
XX	CC	in methods of the invention. Note: The sequence data for this patent is	
XX	CC	not represented in the printed specification but is based on sequence	
XX	CC	information supplied to Derwent by the European Patent Office	
XX	SQ	Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;	
Query Match 27.8%; Score 318.4; DB 6; Length 3387;			
Best Local Similarity 57.8%; Pred. No. 5.8e-76;			
Matches 593; Conservative 0; Mismatches 421; Indels 12; Gaps 1;			
Qy	28	GAAGGGGAAATTTTCAACAGAGAGAGAAAGTCAACAAGTACTCTTTGGAGGTTTAAGAGTGCT	87
Db	2323	GAACTGAATCTTTTCAGTACCTCAACGAAAGACTGTAAGAGGTGGCTCCGTTGCTCC	2382
Qy	88	CTCTATGTCTACTCTCTCTTTAGCTGGAGAGACAATAGCCACTCTCTTTAGGTAGACTTTAC	147
Db	2383	ATATAGCAATCTTTGTTCATCTTTCTGCCACCACTTGTCTACAGTTCTGGGTAGACTGTAC	2442

Db 723 AGCAGTACCGAGTGGCCATTTACCAGTCTCTAGTGGAGTGTGTAGTCTCATAGG 782
QY 738 ACTTTTGGAAAGTGGTGGGAAAGAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGG 797
Db 783 ACTTTTGGCAAGTGGAGAGTGGGAAACTTTGCCAAGTGAGATGAGAACTACAAATCGG 842
QY 798 GAAAGCTCATACATTTTGTATAAACAATCGGTTCAACGATATCATGCAAGCTTTGTTGAT 857
Db 843 GAAAGTGTATATGTTTGGACTTTAGCTTCGGCAGCTATTTCTCGCAAGTCTACACTCT 902
QY 858 TGGAGTGTGGTTTGATATCGAAGTTTCATCGCTTTTTCCTCAATGTATAGCACTCT 917
Db 903 TGGTCTGTGGGAGTATCTTCAGTCACTCTCTGTGTCTTCCCAATTCATACAGCTGT 962
QY 918 TTGTTTACCAGTGTGCTCTTCTGTCTGTGTCTTCTTCCGATGAGATGAGTGAAT 977
Db 963 GGAATGCCATATAGTTCCAGTGGGAGTATAGTTTCCATGATAGATGAGCAATC 1022
QY 978 CAAGTTGGTTGCAATGTTTGGCCATCTGGGAGTTGTTTCTTATGTTTATCAGCAATTA 1037
Db 1023 AAAAATCTTCTCCATATATTTAGCTATCTCGGCTTCTTTCATTCGTCTATCAGCACTA 1082
QY 1038 TGTCAATGATAGAAAG 1053
Db 1083 CCTCGAGAAAGAAAG 1098

RESULT 7

AAA97925
ID AAA97925 standard; DNA; 1071 BP.
XX
AC AAA97925;
XX
DT 19-JAN-2001 (first entry)
XX
DE A. thaliana PUP1 DNA #7.
XX
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.
XX
OS Arabidopsis thaliana.
XX
PN DE19907209-Al.
XX
PD 24-AUG-2000.
XX
PF 19-FEB-1999; 99DE-01007209.
XX
PR 19-FEB-1999; 99DE-01007209.
XX
PA (FROM/) FROMMER W.
XX
PI Gillissen B, Buerkle L, Andre B, Frommer WB;
XX
DR WPI; 2000-566202/53.
XX
XX Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transport, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.
XX
PS Claim 2; Page 15; 24pp; German.
XX
CC This invention describes a novel nucleic acid encoding a plant nucleobase
CC transporter (I). (I) is produced by complementation of a nucleobase
CC transport (NBT)-defective host cell with a plant gene bank by selection
CC of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
CC expression of (II) (when in antisense orientation), and to produce
CC transgenic crop plants. The transgenic plants have modified nucleobase
CC transport properties, e.g. altered affinity and substrate specificity
CC that may result in more efficient nucleobase transport in leaves, changes

CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX
SQ Sequence 1071 BP; 307 A; 236 C; 208 G; 320 T; 0 U; 0 Other;
Query Match 27.0%; Score 308.6; DB 3; Length 1071;
Best Local Similarity 58.3%; Pred. No. 1.7e-73;
Matches 589; Conservative 0; Mismatches 404; Indels 18; Gaps 2;
QY 111 TGGAGAGCAATAGGACCTCTCTTAGGTAGACTTTTACTACGAAAAGCGGTAAAGACAC 170
Db 51 TGGCAATCAGTGTGTACAAATCTGGGCAGACTATATGAAATGGAGGAAACACAA 110
QY 171 ATGGCTCGAAACCTTGGTTTTCAGCTTGTAGGGTTTCTTTAAACCTTCTTGTATATTA 230
Db 111 ATGGCTAGCAACGGTAGTTTACGTTGTAGGCTTCTTATCTTACATCATCTCTT 170
QY 231 CTTAAAGCCTGAGCGCTCAAGACTTAAACCAATACCAAAAAAATCTATCTTCTCTT 290
Db 171 GTCAAGTCAAAACACATACAACTCAGAGAGATGCCAAATTA-----CCTCACTTAG 224
QY 291 GACACTATCTTTAGTATATTTGGACTTGGCTTGGCTTGTCTGGACATTTGTTGTA 350
Db 225 GAAACCGTCATTAGTTTACATAGTCTTGGACTTCTTGTAGGAGCAGCTTGTACCTATA 284
QY 351 CTCATTTGGGCTACTTTTACCTTCTCTGCTCAACTTTTCTTTTGATCTCTGCGTCGCAAT 410
Db 285 TTCCATTTGGACTGCTTTTACCTTCTTACCTTCTTACCTTCTTCCCTGATCTGTCATCAGAT 344
QY 411 GGCCTTTAAACGCGCTTCTCTTCTTCTTAACTCACAATAAATCACAACCAATTTATCT 470
Db 345 AGCCTTCAACGCTTCTCTTCTTCTTCTTCTTCACTCACAATAAATCTTCTCTATCATTTT 404
QY 471 CAATTTCACTTCTTCTTAAACCAATCTTCTACACTTCTTGTATATCAACATGAACACAA 530
Db 405 GAAATTTCTTTTCT 462
QY 531 ATCTCCCT 590
Db 463 -----TCAGATTTCCAAAAAGTTTACAAAAGAGAGTATGTCAAAGGTTTCTGATG 512
QY 591 CGCGTCCGTAGCTCAGCTGGTTTATCTCTGGTGGCTTCTTTTAAACAGATTAACCGTTCA 650
Db 513 CACCGTGGTGATCTCTGGGTTTGGTCTACTCTTATCCCTTACAACAGCTAGCCCTTTCG 572
QY 651 AAAGATTCTAAAGAAATACACATTTCAAGCTATTTTAGACATGGCCACATATCCGCTCTAT 710
Db 573 TAAAGTTTTTAAAGAAAGCAAACTTTCTCAGAAAGTTTATTAATATGATAATCTACATGAGTCT 632
QY 711 GGTAGCTACTTGTGTAGTTGTGTAGGACTTTTTCGAAAGTGGTGGGAAAAAGCTGAG 770
Db 633 AGTGCCAGTTGTGTAGGTTGGTGGGCTTTTCTAGTAGCAGTAGGAAAACTTTGAG 692
QY 771 TACAGAAATGGAAGAGTTTCAACTAGGAAAAAGCTCATACATTTTGTATAAAACATCGGTTTC 830
Db 693 CAGTCAATGGAAGAACTACAACTTTGGGAAGGTATCTCTATGTCATGAACCTAGTGGAC 752
QY 831 AACGATATATGCGCAAGCTTGTTCATTTGGAAGTGTGGTTGTATATGCAAGATTTTCATC 890
Db 753 AGCTGTACCTGGCAGGTATTTCTCCATCGTTTGCAGAGACTGATCTTCGAGCTTTTCTTC 812
QY 891 GCTTTTTTCCAAATGTACAAAGCACTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 950
Db 813 CCTATTTCAAAATGCAATAGCGCTTTGGGACTCCCGGTGTTCTATCTCTGGCTGTCTAT 872
QY 951 CTTCTCCGTGATGAGATGAGTGGAAATCAAGTTGGTTGCAATGTTTTTGGCCATCTGGGG 1010
Db 873 CATTTTCCATGACAAAAATGAACGGCTTAAAGGTGATTTCTATGATTTCTAGCTATTTGGGG 932
QY 1011 ATTTGTTTCTTATGTTTATCAGCAATATGTCATATGATAGAAAGCCAGAAAGACCAAGA 1070
Db 933 TTTCTGATCTCTATGTCTACCAACAATATCTTGTATGAAACAACTTTGAAAGAAAGTAATGA 992


```
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 26.6%; Score 304.8; DB 3; Length 2660;

Best Local Similarity 58.1%; Pred. No. 2.6e-72;

Matches 587; Conservative 0; Mismatches 402; Indels 21; Gaps 2;

```
QY 70 TGGAGGTTAAGAGTGTCTCTATGTCACTCTCTCTTCTAGCTGGAGAGACAATAGCCACT 129
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 TGGTGGTCTTGTATCCGTAAACATCTTTTCCCTATTTGGTGGTCAAGCTCTTCGGTT 250
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 CTCCTAGGTAGACTTTACTACGAAAGGCGGTAAAGCACATGCTCGAAACCTTGGTT 189
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 251 CTTCTTGAGTGTCTTTACTATGATGAAGGTGGAAACAGTAATGATGGCAACTCTTGT 310
QY 190 CAGCTTGTAGGGTTTCCTTTAAACCCTTCCTTGTCTATTATTACTTAAAGCCTGAGCGCTCC 249
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 311 CAACGGGTCTTTTCCCTATACT-----CTATATCCCGCTTTTGTCTACTTCGGTCT 361
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 250 AAGACTAAACCATTAACAAAAAACTACTTCTTCTTCTTTGACACTATCTTTAGTGTAT 309
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 TCGGCTAGTAGAGTCTTCGGAGAGTTCTGTTCTCACTCAAGTACATATGTTTGTATCTAT 421
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 310 ATTGACATTGGCTTGTGTTGCTGGACATTTGTTATTTTGTACTCATTTGGGGTACTTTTAC 369
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 GTTTTGTCTGGTGTGATCAITGCTGGAGATAATATGTTATATACTCTCTGTGTGGACTTTTGTAC 481
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 370 CTTCTGTCTCAACTTCTCTTTTGATCTCTGCTCGCAATTTGGCTTTTAAAGCCCTCTTTC 429
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 CTCCTGTGATCGACGTATTTCGCTCATTTTGGCTACTCAGTTAGCTTTTCAACGGGTGTTC 541
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 430 TCTTACTTCTTAAACTCAAAAAATCAACCATTTATATCTCAATTCACCTTGTCTCTTAA 489
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 542 TCTTATTTTCATCAATGTCTCAGAAGTTCACTGCTTTGATTCTCAACTCCGTTGTTCTCTTGG 601
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 490 ACCATATCTTCTACACTTCTTTGTTATCCAAATGNAACCAGAAATCTCCCTCTTCTACTTCA 549
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 TCGTTCTCCGCTGCTTTGATAGCTCTCAATGATGATGCGGATCTCTCTTCTGTT----- 655
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 550 AAGTCGCGAGCAAGTCCCAAGTATGTGATTCGATACATCTCGCGGTCTGGTAGCTCAGCT 609
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 656 -----GTCTCAAGTCTAAGTATATTTGTTGGTTTGTGTACACTTGTCTGCTGTCT 709
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 610 GGTTATCTCTGTGCTTTCTTTAAACAGATTACCGGTTTCGAAAAGATTCTAAAGAAATAC 669
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 710 CTCATTCTCTGTGTCTATCTCTTATGCACTTCGTTTCGAGAGATTCTGAAGAGGAG 769
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 670 ACATTCAAGGCTATTTTAGACATGGCCACATATCGCTCTATGTTAGTACTTGTGTAGTT 729
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 770 ACGTTTCTGTGTTCTTTGAAATGCAATCTACACTTCTTTAGTGGCGACTTGTGTTTCG 829
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 730 GTGTAGGACTTTTGGAAAGTGGTGGTGGAAAGCTGAGTACAGAAATGGAAGAGTTT 789
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 830 GTTATAGGGCTTTTGTAGCGGGGAATGGAGACGCTGCATGGGGAATGGAAGGTTAT 889
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 790 CAACTAGGGAAGCTCATACATTTTGATAAACATCGGTTTCAACGATATCATGGAAGCT 849
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 890 CATAAAGGGCAAGCTCTTATGTACTGACCTTGTCTGACAGCAGTACTTGGCAAGTG 949
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 850 TGTGTTGATTGGAAGTGTGTTGATTATCGAAGTTTTCATCGCTTTTTCCTCAATGTCTATA 909
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 950 TGTCTGTGGAGTCGTGGGTTTGATATTTCTGGTGAAGCTGCTCTTCTCAAAACGCTCATT 1009
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 910 AGCACTCTTCTTTACCAAGTTGTGCTGTTCTTCTGTTGTGCTTCTTCCGTGATGAGATG 969
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1010 AGTAGGCTCTCTAGCTGTGACTCCACTAGCAGCTTTGGTTGTGTTCCGTGATAAATG 1069
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 970 AGTGGAAATCAAGTTGGTTGCAATGTTTGTGGCCATCTGGGGAATTTGTTTCTTATGTTTAT 1029
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1070 AGTGGTTAAGATTATGCGCAATGCTGATCCCTATTTGGGGTTTCGCTTCTTATGTTTAC 1129
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1030 CAGCATTTATGTAATGATAGAAAGCCAGAGAACAGACCAAGAGCTTCTCTCA 1079
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1130 CAGAAATCATATTGATGACTTGAAGAGTAAGACAGCAGCACCAACAGCTCA 1179
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 9

ADG88190

ID ADG88190 standard; cDNA; 1047 BP.

XX AC ADG88190;

XX AC

DT 22-APR-2004 (first entry)

XX

DE A. thaliana RPP4-upregulated pathogen infection-related gene #632.

XX

KW Pathogen infection-related gene; plant; Peronospora parasitica;
 KW defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;
 KW fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
 XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX WO200222675-A2.

PN PN

XX 21-MAR-2002.

PD 21-MAR-2002.

XX 14-SEP-2001; 2001WO-US028506.

PF 14-SEP-2001; 2001WO-US028506.

XX 15-SEP-2000; 2000US-0232778P.

PR 15-SEP-2000; 2000US-0232778P.

XX 22-JUN-2001; 2001US-0300183P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

PA (UNYC-) UNIV NORTH CAROLINA.

PA (GLAZ/) GLAZEBROOK J.

PA (WANG/) WANG X.

PA (DANG/) DANG J L.

PA (EULG/) EULGEM T.

PA (ZHUT/) ZHU T.

XX Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;

PI WPI; 2002-292409/33.

XX Novel isolated polynucleotide, useful for conveying pathogen resistance

DR to plants, and for identifying plants infected with a pathogen.

XX Claim 3; SEQ ID NO 632; 605pp; English.

XX The invention relates to 691 Arabidopsis thaliana genes (ADG87559--

CC ADG87557) whose expression is altered in response to pathogen infection,

CC and to homologues of these genes from other plants or fungi, especially

CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),

CC cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The

CC expression of genes of the invention was upregulated or downregulated in

CC Arabidopsis plants infected with the oomycete Peronospora parasitica,

CC indicating that they play a role in defence mechanisms. The genes of the

CC invention are regulated by RPP7 or RPP8 which act via unconventional

CC signalling cascades, or by the RPP4-dependent pathway. The invention also

CC relates to polypeptides encoded by the pathogen infection-related genes;

CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)

CC ; expression cassettes, host cells and pathogen-resistant transgenic

CC plants and their progeny comprising a polynucleotide of the invention;

CC and a method of identifying a plant cell infected with a pathogen. The

CC polynucleotide sequences and methods of the invention are useful for

CC identifying plants infected with a pathogen, and for conferring

CC resistance to pathogens such as oomycetes, fungi, bacteria, viruses,

CC nematodes and insects (e.g., aphids). The present sequence represents an

CC Arabidopsis thaliana gene whose expression is altered in response to

CC Peronospora parasitica infection. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1047 BP; 309 A; 225 C; 199 G; 314 T; 0 U; 0 Other;

XX Query Match 26.1%; Score 298.4; DB 6; Length 1047;

XX Best Local Similarity 58.3%; Pred. No. 1e-70;

XX Matches 571; Conservative 0; Mismatches 391; Indels 18; Gaps 2;

QY 140 GACTTTACTAGCAAAAGCGGTAAAGACATGCTCGAAACCTTGTTGAGTTGAG 199

Db 71 GAGTATCTATGACACGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCACCTTGTG 130

QY 200 GGTTCCTTTAACCTTCTGCTATTATTACTTAAGCCTGAGCGGTCCAGACTATAA 259

Db 131 GCTTTCCTGCTGCTACTTCCATATTATATCTTGTCATTTAAACACATGCAACAACCTGATA 190

QY 260 CCATTACCAAAAAAATACTCTTCTCTTGGACACTATCTTTAGTGATATTGGACTTG 319

Db 191 GAGATGAAAAAGAAACCTCACCTAG-----GAACCGTGATTGGTTTACGTAGTGCTTG 244

QY 320 GCTTGCTTTGTTGCTGGACATTTGATTGTTGACTCATTTTGGGCTACTTTTACTCTTCTCT 379

Db 245 GACTTCTTTGAGGACAGATTGCTATCTGTACTCCATTGGAGCTTCTTTTACTTACCCGTTT 304

QY 380 CAACTTCTCTTTGATCTCTGCGTGCCTTAAACGCGCTTTTAAACGCGCTCTTCTTCTTACTTC 439

Db 305 CTACCTATTCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 364

QY 440 TAAACTCAGAAAAAATCAGACCAATTTATATCTCAATTTCAATTTCAATTTCAATTTCAATTT 499

Db 365 TTAACCTCAGAAAAAATTTACCCCTTATCATTTTAAATTTCTTTTCTCTCTCTCTCTCTCTCT 424

QY 500 CTACACTTCTTTGTTATCCACATGACCAAGTCTCCCTCTTCTTCTTCTTCTTCTTCTTCTTCT 559

Db 425 CCACCTTCTTGTGCTTCAATTAATGAGGAGACAGACTCCACAAAGTTTACAAAAAGGAG- 481

QY 560 CCAAGTCCAAGTATGTTGATTTGGATACATCTGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 619

Db 482 -----AGTATGTCMAAGGTTTCATATGCAACGTTGCTGCTGCTGCTGCTGCTGCTGCTG 532

QY 620 TGGTCTTTCTTTTAAACAGATTACGCGTTTCAAAAGATTCTAAAGAAATACACATTTCAAGG 679

Db 533 TAGTCTTTATCCCTACAAACAGCTAGCCTTTCTTAAAGTCTTAAAGAGCAAAATTTCTCAG 592

QY 680 CTATTTTAGACATGGCCACATATCCGTTCTATGTTAGTACTTGTGTAGTTGTGTGTAGGAC 739

Db 593 AAGTTATGGATATGATAATCTACGTGAGTCTAGTGGCCAGTTGTGTTAGCGTGGTGGGGC 652

QY 740 TTTTGGAGTGGTGGGTGAAAAAGCTGAGTACAGAAATGGAAGAGTCTTCACTAGGGA 799

Db 653 TTTTGGTACAGTGGTGGGAAACTTTTGGAGCTGAAATGGATACTACAAACATGGGA 712

QY 800 AAAGCTCATACATTTTGATAAACATCGTTTCAACGATATCATGGCAAGCTTGTGTTGATTG 859

Db 713 AGGTATCCTACATTTATGAACCTAGTGTGGACAGCTGTGTACCTGGCAGGTATTTCTCCATCG 772

QY 860 GAAGTGTGTTGATTTGATTCGAAGTTTCATCGCTTTTTCATGTCATAGCACTCTTT 919

Db 773 GTGGCACAGGACTGATCTTCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 832

QY 920 GTTTTACCAGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979

Db 833 GACTCCAGTGGTTCCTATCTTGGCTGTAATCATTTTCCATGACAAATGATGGGTAA 892

QY 980 AGTTGGTTGCAATGTTTTCGCCCATCTGGGGAATTTGTTTCTTATGTTATCAGCATTTATG 1039

Db 893 AGGTGATTTCTATGATCCTAGCTATTGCGGTCTTCACTTCTCTATGCTTACCAACAATATC 952

QY 1040 TCATGATAGAAAGCCAGAGAGACCAAGAGCTTCTCTAGTCTTAAAGAGAGAGAGAAC 1099

Db 953 TTGATGACAAAAAATCTTGAGAAAAAATCATGAAATCAACAACACAGAAATCCCTGACCCAC 1012

QY 1100 AAAAACAAGTAGATACCAATT 1119

Db 1013 CAGAGCAGAGAGAGTCAACT 1032

RESULT 10

AAA97924

ID AAA97924 standard; DNA; 1081 BP.

XX AAA97924;

XX 19-JAN-2001 (first entry)

XX A. thaliana PUP1 DNA #6.

XX PUP1; transgenic plant; nucleobase transporter; apical dominance;

KW flowering behaviour; senescence; pesticide distribution; ds.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX DE19907209-A1.
XX 24-AUG-2000.
XX 19-FEB-1999; 99DE-01007209.
XX 19-FEB-1999; 99DE-01007209.
XX (FROM/) FROMMER W.
XX Gillissen B, Buerkle L, Andre B, Frommer WB;
XX WPI; 2000-566202/53.
XX Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transport, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.
XX Claim 2; Page 14-15; 24pp; German.
XX This invention describes a novel nucleic acid encoding a plant nucleobase
CC transporter (I). (I) is produced by complementation of a nucleobase
CC transport (NBT)-defective host cell with a plant gene bank by selection
CC of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
CC expression of (II) (when in antisense orientation), and to produce
CC transgenic crop plants. The transgenic plants have modified nucleobase
CC transport properties, e.g. altered affinity and substrate specificity
CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX
XX Sequence 1081 BP; 322 A; 231 C; 205 G; 323 T; 0 U; 0 Other;

Query Match 26.1%; Score 298.4; DB 3; Length 1081;
Best Local Similarity 58.3%; Pred. No. 1e-70;
Matches 571; Conservative 0; Mismatches 391; Indels 18; Gaps 2;

QY 140 GACTTTACTACGAAAGCGGTAAAGACATATGCTGCAACCTTGGTTCAGCTGTAG 199
Db 89 GAGTATACTATGACACCGGAGAAACAGTAAATGGCTGACCAACGGTAGTCAACTTTG 148
QY 200 GGTTCCTTTAAACCTTCCTTGTATTTACTTAAAGCCTGAGCGCTCAAGACTAAA 259
Db 149 GCTTTCCTGTGCTACTTCCATATATATCTTGTCAATTTAAACACATGCAACACTGATA 208
QY 260 CCATTACCAAAAAAATCTCTTCTTCTTGACACTATCTTTAGTGATATTTGGACTTG 319
Db 209 GAGATGGAAGAAAGAACCTCACCTAG-----GAACCGTGTATTGGTTACGTAGTCTTG 262
QY 320 GCTTGTCTTGTGTGACATATGTTATTTGTACTCAATTTGGGCTACTTTACCTTCTGTCT 379
Db 263 GACTTCTTGTAGGACAGATGCTATCTACTCCATTTGGACTTCTTTACTTACCCGTTT 322
QY 380 CAACCTTCTCTTTGATCTCTGGTGGCAATGGCTTTTAAACCGCTTCTCTACTTCC 439
Db 323 CTACCTATTCCCTGTATCTGTGATCTCAGTTAGCCTTCAATGCTTCTTCTCTATTTC 382
QY 440 TAAACTCAGAAAAAATCAGACCACTTTATATCTCAATTCATCTTCTCTTAAACCATATCTT 499
Db 383 TTAACCTCAGAAAAAATCAGACCACTTTATATCTTAAATCTCTTCTCTCTACTATATCTT 442
QY 500 CTACACTTCTTGTATTCAGACATGAAACAGAAATCTCCCTCTCTTCTACTTCAAAAGTCGACG 559
Db 443 CCACCTTACTTGCATTTCAATAATAGGAGACAGACTCCACAAAAAGTTACAAAAAGGAG--- 499
QY 560 CCAAGTCCAAGTATGATGATGATATCATCTGCGGCTCGGTAGCTCGCTGTTATCTC 619
Db 500 -----AGTATGTCAAGAGTTTCATATGACCGGTCTGCTGCTGCTGTTATGCTC 550

QY 620 TGGTGCTTTCTTTAA CAGATTACGGTTTCGAAAGATTCTTAAGAAATA CACATTCAAGG 679
Db 551 TAGTCTTATCCCTACAA CAGCTAGCTTTCTAAAGAGTCTTAAGAGCAAAATTTCTCAG 610
QY 680 CTATTTTACACATGSCCACA TATCCGTCTATGGTAGCTACTTGTGTAGTTGTGGTAGGAC 739
Db 611 AAGTTATGGATATGATNA TCTACGTGAGTCTAGTGGCCAGTTGTGTAGCGGTGGTGGGC 670
QY 740 TTTTGGAAAGTGGTGGTGGAAAAAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGGGA 799
Db 671 TTTTGTCTAGCAGTGAAGTGGAAAACTTTGAGCAGTGAATGGATAACTACAAACATGGGA 730
QY 800 AAAGCTCATACATTTTGTATAAACA TCGGTTCAACGATATCATGCGAAGCTTGTGTTGATTG 859
Db 731 AGGTATCTCATATTATGAACCTAGTGTGACAGCTGTTACCTGGCAGGTATTCTCCATCG 790
QY 860 GNAAGTGTGGTTTGGATTATCGAAGTTTCATCGCTTTTTCATGTCATAGCACCTCTTT 919
Db 791 GTGGCACAGGACTGATCTTCGAGCTCTCTCTATCTCAATCAATAAGCGTTTTGG 850
QY 920 GTTTACCAAGTTGTGCTGCTTCTTCTTCTCGTGATGAGATGAGTGAATCA 979
Db 851 GACTCCCAAGTGGTTCCTATCTTGGCTGTAAATCAITTTTCATGACAAAATGAATGGTTAA 910
QY 980 AGTTGGTTTCAATGTTTGGCCATCTCGGGAATTTGTTTCTTATGGTTATCAGCATTATG 1039
Db 911 AGGTGATTTCTATGATCTCTAGCTATTTGGGGTTTTCACCTTCTCTATCTCTACCAACAATATC 970
QY 1040 TCAATGATAGAAAGCCAGAAAGAACCAAGACAGCTTCTCAGTCTTAAAGAGAGAAAGAAC 1099
Db 971 TTGATGACAAAAAATCTTGAAGAAAAATCATGAATCACAACACAGAAATCCCTGACCCAC 1030
QY 1100 AAAAAACAAGTAGATACCATT 1119
Db 1031 CAGAAGCAGAGAGTCAACT 1050

RESULT 11
AAA97922
ID AAA97922 standard; DNA; 1293 BP.
XX
AC AAA97922;
XX
DT 19-JAN-2001 (first entry)
XX
DE A. thaliana PUP1 DNA #4.
XX
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.
XX
OS Arabidopsis thaliana.
XX
PN DE19907209-A1.
XX
PD 24-AUG-2000.
XX
PF 19-FEB-1999; 99DE-01007209.
XX
PR 19-FEB-1999; 99DE-01007209.
XX
PA (FROM/) FROMMER W.
XX
PI Gillissen B, Buerkle L, Andre B, Frommer WB;
XX
WPI; 2000-566202/53.
XX
PT Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transport, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.
XX
PS Claim 1f; Page 13; 24pp; German.
XX
CC This invention describes a novel nucleic acid encoding a plant nucleobase

CC transporter (I). (I) is produced by complementation of a nucleobase
CC transport (NBT)-defective host cell with a plant gene bank by selection
CC of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (ii) in prokaryotic or eukaryotic cells, for inhibiting
CC expression of (ii) (when in antisense orientation), and to produce
CC transgenic crop plants. The transgenic plants have modified nucleobase
CC transport properties, e.g. altered affinity and substrate specificity
CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX
SQ Sequence 1293 BP; 345 A; 288 C; 247 G; 413 T; 0 U; 0 Other;

Query Match 25.7%; Score 294.4; DB 3; Length 1293;

Best Local Similarity 56.3%; Pred. No. 1.4e-69;

Matches 578; Conservative 0; Mismatches 436; Indels 12; Gaps 1;

```
QY 28 GAAGGGAATTTTCAACAGAGAGAGAGTACAGTCTCTTGAGGTTAAGAGTGCT 87
DB 181 GAAATCGAATCTTCGTGCTGCTCAATCGAAGAACTATAAGAAATGGCTTCGTATTTCC 240
QY 88 CTCTATGTCACCTCTCTTCTAGCTGGAGAGACAATAGCCACTCTCTTAGGTAGACTTTAC 147
DB 241 ATTTACGTGCTCTTCTGCTGCTGCCAAGCACTTCTACAAATTTGGGCAGAGTTAC 300
QY 148 TACGAAAAGGGGGTAAAGCAATCGCTCGAAACCTTGGTTCAGCTGTAGGTTTCC 207
DB 301 TATGAAAATGGTGGGAAGAGTACATGATGGGAACACTTGTCCAACTAATGGCTTCCCT 360
QY 208 TTAACCCCTTCTGTATATTTACTTAAAGCTGAGCGTCCAGACTTAAACCAATACC 267
DB 361 GTTCTGTTCTCTTCCGCTCTCTTTCCCAACCAAAATCCCAACCAAGAGACAGAT 420
QY 268 AAAAAAATCTCTTCTCTTCTGACACTATCTTTAGTGTATATTTGGACTTTGGCTTGT 327
DB 421 TTCCAGAAATCTCTCTCTCTTCCACCAATCTTGATCAGTTTACATCGTTACTGGACTATTA 480
QY 328 GTTGTGACATGTATTTGTATCTATTTGGGCTACTTTACCTTCCCTGCTCAACTTTC 387
DB 481 GTGTCTGCTAATCTATATGTCTCTCTGTTGTTTACTATATCTTACCAGTTTCTACTTTC 540
QY 388 TCTTTGATCTGCGCTCGCAATGGCTTTTAAAGCGGCTTCTCTTACTTCTCTAAACTCA 447
DB 541 TCCTCATCTTGGCTTCAAAATGGCTTCACTGCTCTCTCTCATATTTCTTAAACTCG 600
QY 448 CAAAAAATCACACCAATTTATCTCAATTCACCTTGTCTCTTAAACCATATCTTCTACACTT 507
DB 601 CAGAAGTTTCAACACCTTTCATTTGTGAATCTCTGTTTCTTCTTACTATTTCTCTGCCCTC 660
QY 508 CTTGTTATCCAAATGAAACCGAATCTCCCTCTTCTTCAAGTCCGACGCAAGTCC 567
DB 661 CTCGTGGTCAACACTGATTCGGAA-----AACACAGCAAAAGTGTCTAGAGTA 708
QY 568 AAGTATGTATGGATACATCTGCGGCTCGTAGCTCAGCTGGTTATTTCTCTGGTCTT 627
DB 709 AAATATGTATAGGAATATATGTACCAATGTGTGCTCTCTGTTGGAATGGATTTGCTGCTA 768
QY 628 TCTTTAACAAGATTACGCGTTTGAAAGATTCTTAAAGAAATACATTTCAAGGCTATTTTA 687
DB 769 TCCCTGGTACAACTGATCCTCAGGAAGGTTTAAAGAGCAACATTTCTCAACGGTCACT 828
QY 688 GACATGGCCACATATCCGTCTATGTAGTACTTGTGTAGTGTGGTGTAGGACTTTTGGGA 747
DB 829 GACTTGTGCTGTACCAATCTCTAGTTCGAAGCTGTGTGTTCTCATAGGACTTTTCGCA 888
QY 748 AGTGTGTGGTGAAGAAAGCTGAGTACAGAAATGGAAGTTCACACTAGGGAAGCTCA 807
DB 889 AGCGGGAGTGGAAACCTTTTAAAGTGAATGGAAACCTCAAACTGGGGAAGTGGCA 948
QY 808 TACATTTTGATAAATCATCGGTTTCAACGATATCATGCGCAAGCTTGTGTTGATTTGGAAGTGT 867
```

```
DB 949 TACCTTATGACTTTTGCGCTCGATAGTATTTCTTGGCAAGTCTACACCATTTGGCGTCTG 1008
QY 868 GGTTTGATTATCGAAGTTTTCATCGCTTTTTCCTCAATGTCATAGCACTCTTTGTTTACCA 927
DB 1009 GGACTGATCTTTGAGTCACTCTTCTGTGTTCTTCCAAATTCCTAATCTGTGGGATTCCT 1068
QY 928 GTTGTGCTGTTCTTCTGTTGTTCTTCTCGTGATCAGATGAGTGAATCAAGTTCGTT 987
DB 1069 ATAGTTCAGTTGTAGCAGTATTTGTTTTCATGATAAATGAACGGGTCAAGATCTTC 1128
QY 988 GCAATGTTTGGCCATCTGGGATTTGTTTCTTATGTTTATCAGCAATATGTCATATGAT 1047
DB 1129 TCCATCATTTAGCTATCTGGGATTCATTTCTTCTATCAGCACTACCTCGACGAA 1188
QY 1048 AGAAAG 1053
DB 1189 AAGAAG 1194
```

RESULT 12

ABN85767

ID ABN85767 standard; cDNA; 83698 BP.

XX AC ABN85767;

DT 21-OCT-2002 (first entry)

DE Arabidopsis yellow stripe-like 4 encoding cDNA SEQ ID NO 9.

KW Maize; transgenic; plant; yellow stripe-like; ysl; Arabidopsis;

KW iron uptake; bioremediation; yellow stripe 1; ysl; gene; ss.

OS Arabidopsis sp.

XX WO200240688-A2.

PD 23-MAY-2002.

XX 16-NOV-2001; 2001WO-US043101.

PR 16-NOV-2000; 2000US-0249222P.

PA (UYVA) UNIV YALE.

XX Walker EL, Dellaporta S;

XX WPI; 2002-490144/52.

XX P-PSDB; ABB83920.

PT New yellow stripe1 and yellow stripe1-like genes, useful for altering the

PT distribution of iron within the plant body so that edible parts of crop

PT plants have more iron, or for producing plants useful in enhancing iron

PT uptake from soil.

XX Claim 1; Page 114-159; 187pp; English.

XX The invention relates to an isolated nucleic acid molecule (I), maize

XX yellow stripe 1 (ysl) or yellow stripe-like (ysl) from Arabidopsis

XX (ABN85763-ABN85771). (I) is useful for generating transgenic plants which

XX can be used for enhancing iron uptake from soil and for bioremediation of

XX metal or heavy metal contaminated soil. (I) may also be used to alter the

XX distribution of iron within the plant body so that edible parts of crop

XX plants have more iron. Transgenic plants may also be used in conventional

XX plant breeding schemes to produce progeny which also contain the gene of

XX interest. The present sequence is that of the Arabidopsis ysl encoding

XX cDNA of the invention

XX SQ Sequence 83698 BP; 26295 A; 15150 C; 15104 G; 27149 T; 0 U; 0 Other;

Query Match 21.2%; Score 243.2; DB 6; Length 83698;

Best Local Similarity 54.4%; Pred. No. 5.9e-55;

Matches 547; Conservative 0; Mismatches 438; Indels 21; Gaps 2;

PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-01511080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-01511303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153707P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145112P.	PR 22-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161922P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0162142P.
PR 04-AUG-1999;	99US-0147302P.		
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147933P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 20-AUG-1999;	99US-0149929P.		
PR 23-AUG-1999;	99US-0149902P.		
PR 23-AUG-1999;	99US-0149930P.		
PR 25-AUG-1999;	99US-0150566P.		

Query Match

20.6%; Score 235.8; DB 3; Length 1403;

Best Local Similarity 54.1%; Pred. No. 1.3e-53;

Matches 545; Conservative 0; Mismatches 432; Indels 30; Gaps 2;

Qy 69	TTGGAGGTTAAGAGTGTCCTCTATGTCACCTCTCTCTTAGCTGAGGAGACAATAGCCAC	128
Db 271	TTGGTGGATTCTTGCTCTTTATAAGTATCTTCTCTCATCTCTGCTCAAGCCATTGCTGT	330
Qy 129	TCCTTAGGTAGACTTTTACTACGAAAGGGGTAAAGCACATGGCTCGAAACCTTGGT	188
Db 331	TCCTTTGGTCGGTTTTTATTACACGAAGGTGGAACAGTAAATGGATCTCTACTCTTGT	390
Qy 189	TCAGCTTGTAGGGTTTCCTTTAAACCCCTTCCTTGCTATTTACTTTAAAGCCTGAGCCGTC	248
Db 391	CCAACTTGTTGGTTTTCCGGATCTCTATCTCCCTCTTTGTTCTCTCTTCCACTC	450
Qy 249	CAAGACTAAACCAATTACCAAAAAAATACTCTCTCTCTTCTTGACACATATCTTTAGTGTA	308
Db 451	TTCTTC-----ATCTTCTTGTTCTTTCAAGACTCTGGTTTGGATTTA	492

QY 664 AATACACATTCAGGCTATTTTACACATGGCCACATATCCGTCTATCGTAGCTACTTGT 723
Db 844 AGGAGACCTCTTCAGTGTGTGAACATGCGATATATACAGCTCTCGTGGCAACATTG 903
QY 724 GTAGTTGTGTAGGACTTTTGGAAAGTGGTGGTGGAAAGAGCTGAGTACAGAAATGGA 783
Db 904 GCTTCTCTGTGTGGTATTTGCAAGTGGTGAATGATGACTTTTACAAGGAGAGATGCAT 963
QY 784 GAGTTTCACTAGGAAAGCTCATACATTTTGTATAAATCATCGGTTTCAACGATATCATGG 843
Db 964 GCATTCAGTCTGGGAAGCTGTCATATGTAATGACACTGCTGTGACGGCTATATCTTGG 1023
QY 844 CAAGCTTGTGATTGGAAGTGTGGTTTGTATTATCGAAGTTTCATCGCTTTTTCATAT 903
Db 1024 CAGATGGATGGAGTAAGATTATAGCTATGCTGATGGCCATTTGGGGATTATGTATAT 1083
QY 904 GTC 906
Db 1084 GGC 1086

RESULT 15

ADA70748
ID ADA70748 standard; DNA; 2175 BP.

AC ADA70748;

DT 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 4071.

EX Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.

OS Oryza sativa.

XX WO200300898-A1.

PN 03-JAN-2003.

PD 22-JUN-2001; 2001WO-IB001105.

PF 22-JUN-2001; 2001WO-IB001105.

PR (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

PS Claim 6; SEQ ID NO 4071; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

SQ Sequence 2175 BP; 363 A; 701 C; 671 G; 439 T; 0 U; 1 Other;

Query Match 14.2%; Score 162.2; DB 8; Length 2175;

Best Local Similarity 50.8%; Pred. No. 1.7e-33;
Matches 460; Conservative 0; Mismatches 418; Indels 27; Gaps 2;

QY 70 TGGAGGTTAAGAGTGTCTCTATGTCACTCTCTCTTAGCTGGAGACAAATAGCCACT 129
Db 136 TGGTGGCGATGGTGGCAGTGGACGTCTTCTTCTCATCGCGGTGACATCGCGGACA 195
QY 130 CTCTTAGGTAGACTTTTACTAGAAAAGGCGGTAAAGACACATGGCTCGAAACCTTGGTT 189
Db 196 CTGTGGGCGAGATACTACTACACCAAGGCGCGCAGCAAGTGATATCGGCGTTCTGTG 255
QY 190 CAGCTTGTAGGGTTCTCTTTAAACCTTCTTGTCTATTATTACTTTAAAGCCTGAGCCGTC 249
Db 256 CGGAGCGCGGCTTCCGATACTGTCTTCAACCTCTTCTTCTCCCTCGAAGTCAACG 315
QY 250 AAGACTAAACCAATTAACAAAACAACTACTTCTTCTTGTGACACTATCTTTAGTGTAT 309
Db 316 TCCTCTCTGCACCAACA-----CTCCATGGCCAGCTCGCGTGATATAC 360
QY 310 ATTGGACTTGGCTTGTGTGGACATGTATTTTGTACTCATTTGGGCTACTTTAC 369
Db 361 ATGCTCTGGGCTCATCATCGCGCGCAGACATGATGTACACTGGTGGCCTCAAGTAC 420
QY 370 CTTCTGTCTCAACTTCTCTTTGATCTCTCGGTGCAATTTGGCTTTTAAACGCGCTCTTC 429
Db 421 CTCCCGCTTGGACCTACTCGCTCATCTGTGCCAGCAGCTCGGTTCAATGTGCTTTC 480
QY 430 TCTTACTTCTTAAACTCACAAAAAATCACACCAATTTATATCAATCACTTGTCTCTTA 489
Db 481 TCATACGTGTCAACTCCAGAGGTCACTCTCTGATATTTCAACTCCGTCGTGCTGCTC 540
QY 490 ACCATATCTTACACTTCTTGTATCCACATGAACCAAGATCTCCCTCTTCTACTTCA 549
Db 541 ACCATGTCCGCTTGGCTCATCGGAGTCAGCAAGAGTCTCAGGGGGTCAACG----- 592
QY 550 AAGTCCGACGCAAGTCCAAAGTATGTATTGGATACATCTGCGGGTGGTAGTCTCAGCT 609
Db 593 ----GCGTCTGGGAGGGAAGTATCTGCTCGTTCGTGTGAGCTGGGGGCGTCTGTC 648
QY 610 GGTATTCTCTGGTGTCTTTCTTTAAACAGATTACGCGTTTGAAGAGATTCTTAAAGAAATAC 669
Db 649 ACCTACTCGTGATCTCTCGCGCTGATGCAGCTCACCTTCGAGACCATCATCAAGAACAC 708
QY 670 ACATTCAAGGCTATTTTAGACATGGCCACATATCCGTCTATGTGTAGTCTTGTGTAGTT 729
Db 709 ACCTTCTCAGCGCTCTCTCAACATGCGAGTCTACACGGCGCTCGTGGGACGCGCGCTCG 768
QY 730 GTGTAGGACTTTTGGAGTGGTGGTGGAAAAAGCTGAGTACAGAAAAATGGAAGAGTTT 789
Db 769 GTGTGGGCTGTTCGCGAGCGGCGAGTGGAGGTGCTGAGGGGGGAGATGAAACGCTTC 828
QY 790 CAACTAGGAAAAAGCTCATACATTTTGAATAAATCATCGTTTCAAGATATCATGCAAGCT 849
Db 829 AGGTGGGGCAGATTCTCTACCTGATACGCTGTGTGGGCGCGCTGTCTGGCAGGTG 888
QY 850 TGTGTGATTGGAAGTGTGGTTTGAATATCGAAGTTTTCATCGCTTTTTCATGTGATTA 909
Db 889 GCCAACATCGGGGTGCTCGGCGCTCACTTTCAGGTGTGCGGCTCTTCTTCAAGCTGATC 948
QY 910 AGCACTCTTTTGTATTACAGTGTGCGCTGTCTTGTCTTCTTCTTCCGTGTAGATG 969
Db 949 AGCAGGTGTCTACTGCGGCTTCCCGTCTTCTCGGCTGTGTGTGTTCCACGACAGATG 1008
QY 970 AGTG 974
Db 1009 AACGG 1013

Search completed: November 1, 2004, 07:19:49
Job time : 608.255 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:44:43 ; Search time 5074.79 Seconds
(without alignments)
10669.751 Million cell updates/sec

Title: US-09-913-767-3
Perfect score: 1145
Sequence: 1 ctgtttcatgtgtgtatgggt.....caagtttaggcaagatcca 1145

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4525729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_srs.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1145	100.0	1145	6	AX033546 Sequence
2	1145	100.0	58427	8	ATF15J5
3	1145	100.0	82697	8	AT79A21
4	1145	100.0	194143	8	ATCHRIV48
5	376.6	32.9	73179	8	AP006383
6	320.6	28.0	1106	8	AF370622
7	318.4	27.8	1165	8	AY096558
8	318.4	27.8	1428	8	AY074546
9	318.4	27.8	3387	6	AX507517
10	318.4	27.8	3387	6	AX652015
11	317.2	27.7	1194	6	AX033548
12	308.6	27.0	1071	6	AX033550
13	304.8	26.6	1251	8	BT003358
14	304.8	26.6	1372	8	AY062782
15	304.8	26.6	115421	8	T12C22
16	298.4	26.1	1047	6	AX412868
17	298.4	26.1	1081	6	AX033549
18	294.4	25.7	1293	6	AX033547
19	281.8	24.6	1476	8	AK071254

20	281.8	24.6	1617	8	AK072660
21	281.8	24.6	175645	8	AP004096
22	277	24.2	110036	8	AC142095
23	243.2	21.2	83698	6	AX540653
24	243.2	21.2	83698	8	AB010072
25	237.4	20.7	1117	8	AY096468
26	237.4	20.7	1333	8	AT74535
27	237.4	20.7	104386	8	ATT32A17
28	237.4	20.7	199362	8	ATCHRIV24
29	235.8	20.6	1403	8	AY087844
30	203.8	17.8	1279	8	BT014359
31	197	17.2	133299	8	OSUN00098
32	197	17.2	145739	8	OSUN00157
33	193.8	16.9	1486	8	AK119483
34	190.2	16.6	1152	6	AX653666
35	167	14.6	733	8	AF083706
36	162.2	14.2	2175	6	AX654201
37	157.2	13.7	721	11	BX664674
38	155.2	13.6	1361	8	AK071680
39	130.6	11.4	1155	6	AX653139
40	130.6	11.4	1599	8	AK060241
41	125.4	11.0	1040	6	AX653021
42	125	10.9	131394	8	AP005679
43	119.8	10.5	1067	6	AX033553
44	112	9.8	1049	6	AX033545
45	99.8	8.7	12909	8	AC079991

ALIGNMENTS

RESULT 1
AX033546
LOCUS AX033546 1145 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 3 from Patent WO049152.
ACCESSION AX033546
VERSION AX033546.1 GI:10280290
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE 1
AUTHORS Andre B., Buerkle, L., Frommer, W.B. and Gillissen, B.
TITLE Nucleic acids that code for a nucleobase transporter
JOURNAL Patent: WO 049152-A 3 24-AUG-2000;

ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;
GILLISSEN BERND (DE)

FEATURES
Location/Qualifiers
source 1..1145

/organism="Arabidopsis thaliana"
/db_xref="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Query Match 100.0%; Score 1145; DB 6; Length 1145;
Best Local Similarity 100.0%; Pred. No. 3.4e-267;
Matches 1145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGTTTCATGTGTTGATGTTAGAGTGTCTCTATGTCACTCTCTCTTCTAGTGGAGAGACA 120
DB 1 CTGTTTCATGTGTTGATGTTAGAGTGTCTCTATGTCACTCTCTCTTCTAGTGGAGAGACA 120
QY 61 AAGTACTCTTGGAGTTAAGAGTGTCTCTATGTCACTCTCTCTTCTAGTGGAGAGACA 120
DB 61 AAGTACTCTTGGAGTTAAGAGTGTCTCTATGTCACTCTCTCTTCTAGTGGAGAGACA 120
QY 121 ATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAGCGGTAAAGCACATGGCTCGAA 180
DB 121 ATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAGCGGTAAAGCACATGGCTCGAA 180
QY 181 ACCTTGGTTCAGCTTGTAGGGTTTCCCTTTAAACCCCTTCCTTGTCTATTATTACTTAAAGCCT 240

Db	181		ACCTTGGTTGAGCTTGAGGTTTCCTTTAACCTTCTCTGCTATTATTACTTTAAAGCT	240
Qy	241		GAGCGTCCAAAGACTAAACCAATTAACCAAAAAAATACTTCTCTCTTTGACACTATCT	300
Db	241		GAGCGTCCAAAGACTAAACCAATTAACCAAAAAAATACTTCTCTCTTTGACACTATCT	300
Qy	301		TTAGTGATATGACACTTGGCTTGGTTGCTGCTGGACATTGTTATTTGTACTCATTTGGG	360
Db	301		TTAGTGATATGACACTTGGCTTGGTTGCTGCTGGACATTGTTATTTGTACTCATTTGGG	360
Qy	361		CTACTTTACCTTCCTGCTCAACTTCTCTTTGATCTCTGGTGCCTAAATGGCTTTTAAAC	420
Db	361		CTACTTTACCTTCCTGCTCAACTTCTCTTTGATCTCTGGTGCCTAAATGGCTTTTAAAC	420
Qy	421		CCCGTCTCTTCTTACTTCTTAACTCAAAAAATCACACATTATATACTCAATTCACTT	480
Db	421		CCCGTCTCTTCTTACTTCTTAACTCAAAAAATCACACATTATATACTCAATTCACTT	480
Qy	481		GTTCTCTTAACCATATCTTCTACACTTCTTGTATCCAGATGAACAGATCTCCCTCT	540
Db	481		GTTCTCTTAACCATATCTTCTACACTTCTTGTATCCAGATGAACAGATCTCCCTCT	540
Qy	541		TCTACTTCAAAAGTCGACGCAAGTCCAAAGTATGTGATTGGATACATCTGGCGGTCGGT	600
Db	541		TCTACTTCAAAAGTCGACGCAAGTCCAAAGTATGTGATTGGATACATCTGGCGGTCGGT	600
Qy	601		AGCTCAGCTGTTTATCTCTGGTGCTTTCTTTTAAACAGATTACGGTTCGAAAAAGATTCTA	660
Db	601		AGCTCAGCTGTTTATCTCTGGTGCTTTCTTTTAAACAGATTACGGTTCGAAAAAGATTCTA	660
Qy	661		AAGAAATACATTCAGGCTATTTAGACATGCCACATATCCGTCATGGTAGCTACT	720
Db	661		AAGAAATACATTCAGGCTATTTAGACATGCCACATATCCGTCATGGTAGCTACT	720
Qy	721		TGTGTAGTTGTTGAGGACTTTTGGAAAGTGGTGGTGGAAAAAGCTGAGTACAGAAATG	780
Db	721		TGTGTAGTTGTTGAGGACTTTTGGAAAGTGGTGGTGGAAAAAGCTGAGTACAGAAATG	780
Qy	781		GAAGAGTTTCAACTAGGAAAAAGCTATACATTTTGATATAACATCGGTTCAACGATATCA	840
Db	781		GAAGAGTTTCAACTAGGAAAAAGCTATACATTTTGATATAACATCGGTTCAACGATATCA	840
Qy	841		TGGCAAGCTTGTGATTGGAAGTGTGGTTGATTATCGAAGTTTCATCGCTTTTTC	900
Db	841		TGGCAAGCTTGTGATTGGAAGTGTGGTTGATTATCGAAGTTTCATCGCTTTTTC	900
Qy	901		AATGTCATAAGCACTCTTTGTTTACCACTTGTGCTGTCTTGTGTTGTTCTTCTTCCT	960
Db	901		AATGTCATAAGCACTCTTTGTTTACCACTTGTGCTGTCTTGTGTTGTTCTTCTTCCT	960
Qy	961		GATGAGATGAGTGAATCAAGTTGGTTGCAATGTTTGGCCATCTGGGGATTTGTTTCT	1020
Db	961		GATGAGATGAGTGAATCAAGTTGGTTGCAATGTTTGGCCATCTGGGGATTTGTTTCT	1020
Qy	1021		TATGGTTATCAGCATTATGTCATATAGAACGACGAGAGACCAAGAGCTTCCCTCAG	1080
Db	1021		TATGGTTATCAGCATTATGTCATATAGAACGACGAGAGACCAAGAGCTTCCCTCAG	1080
Qy	1081		TCTAAAGAAAGAAAGAAACAAAAACAGTAGATACCATTCATGTGCCAGCTTAGGCAAG	1140
Db	1081		TCTAAAGAAAGAAAGAAACAAAAACAGTAGATACCATTCATGTGCCAGCTTAGGCAAG	1140
Qy	1141		ATCCA 1145	
Db	1141		ATCCA 1145	

RESULT 2
ATF15J5
LOCUS
DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone F15J5 (ESSA project).

ACCESSION	AL110123
VERSION	AL110123.1
KEYWORDS	GI:5816989
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	1 Bevan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H.W., Mayer, K.P.X., Lemcke, K. and Schueller, C. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	2 (bases 1 to 58427)
AUTHORS	EU Arabidopsis sequencing project.
TITLE	Submitted (30-MAR-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schueller@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ .
FEATURES	Location/Qualifiers
source	1..58427 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /variety="Columbia" /db_xref="taxon:3702" /chromosome="4"
misc_feature	1..2548 /note="overlap to BAC T6K21; please refer to EMBL:AL021889 for analysis and annotation"
gene	3596..5023 /gene="F15J5.10" complement(join(3596..3646,3736..3801,3895..4020,4116..4281,4725..5023))
gene	/gene="F15J5.10" complement(join(3596..3646,3736..3801,3895..4020,4116..4281,4725..5023))
CDS	/note="Contains Eukaryotic initiation factor 4E signature [DPYCKHIIPEKWEPIICANGKWI] contains EST gb:T21435, AA586111" /codon_start=1 /product="translation initiation factor eIF4E" /protein_id="CAB53645.1" /db_xref="GI:5816990" /db_xref="GOA:O23252" /db_xref="Swiss-Prot:O23252" /translation="MAVEDTPKSVTTEAKPNSIENPIDRYHEGDDAERGETAGGSG DGNVDSSKSGVPESHPLSHSWTFWFDNPAVKSKQTSWSSLRPVTFTSTVEEFLY NNMKHSKLAHGDYCFKHIIEPKWEDPIICANGKWTMTFFPKSKDSKSLYTLALI GEQFDHGEITCGAVVNIQRKQISITWTKNASNAQAQVSIQKQKEFLDYNNSIGFII HEDAKLIDRNKNAVTA" complement(3596..3646) /gene="F15J5.10" /number=1 complement(3647..3735) /gene="F15J5.10" /number=1 complement(3736..3801) /gene="F15J5.10" /number=2 complement(3802..3894) /gene="F15J5.10" /number=2 complement(3895..4020) /gene="F15J5.10" /number=3 complement(4021..4115) /gene="F15J5.10"
exon	
intron	
exon	
intron	
exon	
intron	

Db	55173	GAGCGTCCAGAGCTAAACACCTTACCAAAAAAATACTCTTCTCTTGACACTATCT		55232
Qy	301	TTAGGTATATGGACTTGGCTTGCTTGCTCGACATGATATTTGTACTCATTTGGG		360
Db	55233	TTAGGTATATGGACTTGGCTTGCTTGCTCGACATGATATTTGTACTCATTTGGG		55292
Qy	361	CTACTTTACCTTCCTGCTCGACTTCTCTTTTGATCTCTGCGTGCATTTGGCTTTTAAAC		420
Db	55293	CTACTTTACCTTCCTGCTCGACTTCTCTTTTGATCTCTGCGTGCATTTGGCTTTTAAAC		55352
Qy	421	GCCGCTCTCTTACTTCTTAAATCAAAAAATACACCATTTATATACTCAATTCACCTT		480
Db	55353	GCCGCTCTCTTACTTCTTAAATCAAAAAATACACCATTTATATACTCAATTCACCTT		55412
Qy	481	GTTCTCTTAAACATATCTTTACACTTCTTGTTATCAACAATGAACAGAAATCTCCCTCT		540
Db	55413	GTTCTCTTAAACATATCTTTACACTTCTTGTTATCAACAATGAACAGAAATCTCCCTCT		55472
Qy	541	TCTACTTCAAAAGTCGGAGCCAGTCCAAGTATCTGATTTGGATACATCTCGCGGTCGGT		600
Db	55473	TCTACTTCAAAAGTCGGAGCCAGTCCAAGTATCTGATTTGGATACATCTCGCGGTCGGT		55532
Qy	601	AGCTCAGCTGGTATCTCTCGTCTTCTTTTAAACAGATTACGGCTTCGAAAAATTTCTA		660
Db	55532	AGCTCAGCTGGTATCTCTCGTCTTCTTTTAAACAGATTACGGCTTCGAAAAATTTCTA		55592
Qy	661	AAGAAATACACATTCGAAGCTATTTTAGACATGCCCACATATCCGCTATGGTAGCTACT		720
Db	55593	AAGAAATACACATTCGAAGCTATTTTAGACATGCCCACATATCCGCTATGGTAGCTACT		55652
Qy	721	TGCTAGTGTGGTAGGACTTTTGGAGTGGTGGGTGGAAAAAGCTGAGTACAGAATG		780
Db	55653	TGCTAGTGTGGTAGGACTTTTGGAGTGGTGGGTGGAAAAAGCTGAGTACAGAATG		55712
Qy	781	GAAGAGTTTCAACTAGGAAAAAGCTCATACATTTTGATTAACATCGGTTCAACGATATCA		840
Db	55713	GAAGAGTTTCAACTAGGAAAAAGCTCATACATTTTGATTAACATCGGTTCAACGATATCA		55772
Qy	841	TGGCAAGCTTGTGATTTGGAAGTGTGGTTGATTAATCGAAGTTTCATCGCTTTTTC		900
Db	55773	TGGCAAGCTTGTGATTTGGAAGTGTGGTTGATTAATCGAAGTTTCATCGCTTTTTC		55832
Qy	901	AATGTCATAACACTCTTTGTTTACCAAGTTGCGCTGTTCTTGCTGTGTCTTCTTCGGT		960
Db	55833	AATGTCATAACACTCTTTGTTTACCAAGTTGCGCTGTTCTTGCTGTGTCTTCTTCGGT		55892
Qy	961	GATGAGATGAGTGGAAATCAAGTTGGTTCGCAATGTTTTTGGCCATCTGGGGATTTGTTTCT		1020
Db	55893	GATGAGATGAGTGGAAATCAAGTTGGTTCGCAATGTTTTTGGCCATCTGGGGATTTGTTTCT		55952
Qy	1021	TATGGTTATACAGCATTTATGTCATATAGTAAGCCAGAGAAGACCAAGAGCTTCCTCAG		1080
Db	55953	TATGGTTATACAGCATTTATGTCATATAGTAAGCCAGAGAAGACCAAGAGCTTCCTCAG		56012
Qy	1081	TCTAAAGAGAGAGAGACAAACACAGTACATACCATTCATGTCCTAGCTTAGGCAAG		1140
Db	56013	TCTAAAGAGAGAGAGACAAACACAGTACATACCATTCATGTCCTAGCTTAGGCAAG		56072
Qy	1141	ATCCA 1145		
Db	56073	ATCCA 56077		

RESULT 3
LOCUS ATT9A21 82697 bp DNA linear PLN 23-SEP-1999
DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone T9A21 (ESSA project).
ACCESSION AL021713
VERSION AL021713.1 GI:2832689
KEYWORDS Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 Bevan,M., Murphy,G., Ridley,P., Hudson,S., Bancroft,I., Mewes,H.W., Mayer,K.P.X., Lemcke,K. and Schueller,C.
Unpublished
2 (bases 1 to 82697)
EU Arabidopsis sequencing,project.
Direct Submission
Submitted (22-SEP-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferepitz 18a, D-82152 Martinsried, FRG, E-mail: schueller@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de,Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
Location/Qualifiers
1. 82697
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
402..1585
/genes="T9A21.10"
join(402..832,946..1090,1178..1585)
/genes="T9A21.10"
/note="similarity to SPF1 protein, sweet protein, PIR2:S51529 and WRKY protein family, Petroselinum crispum, MNOS:S72443, MNOS:S72444, MNOS:S72445"
/codon_start=1
/product="DNA binding-like protein"
/protein_id="CAA16788.1"
/db_xref="GI:4038481"
/db_xref="GOA:O49720"
/db_xref="GOA:Q8VWJ2"
/db_xref="Swiss-Prot:Q8VWJ2"
/translation="MSNETRDLYNQYSPSSFSLHEMNLPTSPSSYGNLPSONGFNPSTYFTDCLQSSPAAYESLLQKTFGLSPSSSEVFNSSIDOENRDVTVNDVINGACNE TETRVSPNSSSSEADHPGDSKSRKRELVEEDQISKVGKTKTEVKVKOREPRV SFWTKSEVDHLEDGYRWXYGXAVKNSPYPRIANGNENRSYVRCCTCKNVKRVK RSQDPTVITTTTGGQHNHPITNLKSGSAAAFSADLMTPRSFADHMFRTAAYTNG GSVAALDVGYGQSGSVNSPNSSHQVYHQGGEYELLREIFPSIFPKQEP"
402..832
/genes="T9A21.10"
/number=1
833..945
/genes="T9A21.10"
/number=1
946..1090
/genes="T9A21.10"
/number=2
1091..1177
/genes="T9A21.10"
/number=2
1178..1585
/genes="T9A21.10"
/number=3
4778..6211
/genes="T9A21.20"
join(<4778..4921,5023..5187,5263..5552,5638..5866, 5951..6211)
/genes="T9A21.20"
/note="strong similarity to exopolysaccharuronase, Arabidopsis thaliana, PIR2:S34266"
/codon_start=1
/product="polysaccharuronase-like protein"
/protein_id="CAA16789.1"
/db_xref="GI:2832691"

```
/db_xref="GOA:O49721"
/db_xref="TrEMBL:O49721"
/translation="AFVAAMDKACKSSSSVNLIIPRGFSVGLRFGPCTNVSRLT
VRYKASTDLKSPSGGWQFGWINGLITLGGTDFGOGALAPFNCTSDNCKLLP
TSLKFFVGMRTVVRRISSVNSKFFPHIALVECRDFKGRNLNITAPSDPNTDGHIHRS
SNYFRSHIATGDDCVSGGNSQITITISIKCGPHGHSVGLGRPEKONGVGLV
KDKISGTINGIRIKTWSNPGLSAATMTFENIMNVNTPIIIDOSYCPFSSCISN
VPKSISEIYFNRIKRTSSSLVAVQLHCSRGMPCCKVYLENVHLDLSSSDGGRKQSS
NRGNEAVSSSCRNVRYNGTQIPPPCH"
exon      4778..4921
           /gene="T9A21.20"
           /number=1
intron    4922..5022
           /gene="T9A21.20"
           /number=1
exon      5023..5187
           /gene="T9A21.20"
           /number=2
intron    5188..5262
           /gene="T9A21.20"
           /number=2
exon      5263..5552
           /gene="T9A21.20"
           /number=3
intron    5553..5637
           /gene="T9A21.20"
           /number=3
exon      5638..5866
           /gene="T9A21.20"
           /number=4
intron    5867..5950
           /gene="T9A21.20"
           /number=4
exon      5951..6211
           /gene="T9A21.20"
           /number=5
gene      6969..8087
CDS       join(6969..7148,7191..8087)
           /gene="T9A21.30"
           /note="similarity to W30DMY32F, W25DMY32F"
           /codon_start=1
           /product="putative protein"
           /protein_id="CAA16790.1"
           /db_xref="GI:2832692"
           /db_xref="TrEMBL:O49722"
/translation="WVEPGKFEERSHKYSWRLRVSLYVLLLAGETIALLGLRY
YKGGSTWLETVLQPEPSKTIITKKTSFTLSLVYIGLLVAGHCILYSPGL
LYLPVFTSLISQAFNAVFSYFLNSQKITPTFILNSLVLTISLLVIOHEPES
SSTSKSAKSKYVIGYICAVSGSAGSLVSLTDYAFKILKKTPTKAILDMATYPSM
VATCVVVGLPFGSGWKLSSTEMEELQGLKSYLINIGSTISWQACLIGSVGLIIEV
SSLFSNVI STLCLPVVPLAVVFRDEMSGIKLVAMFLAINGFVSIGYQHYVNDRKE
EDQLPQSKEEEEQKQVDIHVQA"
exon      6969..7148
           /gene="T9A21.30"
           /number=1
intron    7149..7190
           /gene="T9A21.30"
           /number=1
exon      7191..8087
           /gene="T9A21.30"
           /number=2
gene      8612..13959
CDS       join(8612..8838,8884..9860,10782..11155,11210..11915,
           12857..13959)
           /gene="T9A21.50"
           /note="similarity to W30DMY32F, W15DMY32F
           contains EST gb:T42237"
           /codon_start=1
           /product="putative protein"
           /protein_id="CAA16792.1"
           /db_xref="GI:2832694"
```

```
/db_xref="TrEMBL:O49724"
/translation="MEITQVIYVNGKQDASRSVDYLILFANLLFLMFSDDHNNIEANLT
GOEWNNTMEIESSVPOSKYKWKLRISIVVYVYENGKSTWMTGLVLQIGPPVLF
PEFSSQNKPKTEADPRKESFTILGSVYIVTGLLVANSYMSVGLLYLVSTFSL
ILASQATAPFSEFLNSOKFTFPFVNSLFLTITSSALLVNTDSENTAKVRKVI
GLICTAGAGIGLLLSVLQILRKVKKQFTSTVTDLVAYOSLVASCVLIGLFPASG
EWTITSEMENYKLGKVPVMTLASIAISWQYITIGVGLIFESSVFSNITAVGLP
IVPVAVIVFHDKNNAKIFSIILAIWGFISFYQHYLDEKKLTKSHTSFVGDPLLLP
ABEGHTNTHDONLEANLIDHEVWTESSSAVPQENYKRWLRVSIYIVFLPCQPLA
TELRLYHENGNSVYVVTLLQLIGPPVILVFRFRIROPKSTDNTNFSQSPFTTLA
SVYLCGLLVSAYALASALATPAPFSEFLNSOKFTPLIVSSLLITVSSALLVNTDS
ENSTNSRVQYVIGFICTIGASAGIGLLLSLQMLFRKVTIKHTSSAVTDLAIYQLV
ASCVVLI GLFASGEWETLPSEMRNYKLGKSVYLTLASAAISWQYITGLVGLIFESS
SVFSNSITAVGLPIVFAAVIVFHDMDASKIFSIILAIWGFISFYQHYLDEKKLNT
SHTSAGDLHLPEVERGHTNIQSDONLEANLIDHEETSPSVQTKNCKRWLRVSIYAI
FVIFCQPLATVILGYVYENGKSTVVTLLQLIGPPVILVFRFRIROPKSTDNTNFS
QSPFTTLASVYLCGLLVSAYALSAVGLLYLVSTFSLILASQATAPFSEFLNS
QKFTPLIVNSLFLTIVSSALLVNTDSENTNSRVQYVIGFICTIGASAGIGLVLSL
IQLFRKVTIKHTSSAVLDLANYQSLVATCVVLI GLFASGEWRTLPSEMRNYKLGKVS
YILTASAAIFWQYITVGCGLIFESSVFSNITAVGLPIVFPVAVIVFHDMDASK
IFSIIILAIWGLFSFYQHYLDEKKLTKCTQKVEBETQL"
exon      8612..8838
           /gene="T9A21.50"
           /number=1
intron    8839..8883
           /gene="T9A21.50"
           /number=1
exon      8884..9860
           /gene="T9A21.50"
           /number=2
intron    9861..10781
           /gene="T9A21.50"
           /number=2
exon      10782..11155
           /gene="T9A21.50"
           /number=3
intron    11156..11209
           /gene="T9A21.50"
           /number=3
exon      11210..11915
           /gene="T9A21.50"
           /number=4
intron    11916..12856
           /gene="T9A21.50"
           /number=4
exon      12857..13959
           /gene="T9A21.50"

Query Match      100.0%; Score 1145; DB 8; Length 82697;
Best Local Similarity 100.0%; Pred. No. 2.2e-267;
Matches 1145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CTGTTTCATGCTTGTATGTTAGTAACTGAAGGGAATTTTCAACAGAGAGAGAGTCAAC 60
DB      6954  CTGTTTCATGCTTGTATGTTAGTAACTGAAGGGAATTTTCAACAGAGAGAGAGTCAAC 7013

QY      61  AAGTACTCTTGGAGGTTAAGAGTGCTCTCTATGTACACTCTCTCTTACTCTGAGAGACA 120
DB      7014  AAGTACTCTTGGAGGTTAAGAGTGCTCTCTATGTACACTCTCTCTTACTCTGAGAGACA 7073

QY      121  ATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAAAGCGGTAAAGACACATGGCTCGAA 180
DB      7074  ATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAAAGCGGTAAAGACACATGGCTCGAA 7133

QY      181  ACCTTGGTTCACCTTGTAGGGTTTCCTTTAACCCCTTCCTTGTATATATTACTTAAGCCCT 240
DB      7134  ACCTTGGTTCACCTTGTAGGGTTTCCTTTAACCCCTTCCTTGTATATATTACTTAAGCCCT 7193

QY      241  GAGCCGTCCAAGACTAAACACCATTAACCAAAAAAATCTCTTCTTCTTGTACACTATCT 300
DB      7194  GAGCCGTCCAAGACTAAACACCATTAACCAAAAAAATCTCTTCTTCTTGTACACTATCT 7253

QY      301  TTAGTGTATATTGGAGCTTGGCTTGCTTGTGTGGACATTGTATTTTGTACTATTGGG 360
```

Db	7254	TTAGTGTATATTGGACTTGGCTTGCTGTGTGCGACATGTATTGTGACTCAATTGGG	7313
QY	361	CTACTTTACCTTCCTGCTCAACTTTCTCTTTGATCTCTGCGTCGCAATTGGCTTTTAAC	420
Db	7314	CTACTTTACCTTCCTGCTCAACTTTCTCTTTGATCTCTGCGTCGCAATTGGCTTTTAAC	7373
QY	421	GCGCTCTCTTCTTACTTCTTAACCTCAACAAAAATCACACCACTTTATATCTCAATTCACTT	480
Db	7374	GCGCTCTCTTCTTACTTCTTAACCTCAACAAAAATCACACCACTTTATATCTCAATTCACTT	7433
QY	481	GTTCTCTTAACCATATCTTCTACACTTCTTGTTATCCAAATGACCAAGATCTCCCTCT	540
Db	7434	GTTCTCTTAACCATATCTTCTACACTTCTTGTTATCCAAATGACCAAGATCTCCCTCT	7493
QY	541	TCCTACTTCAAGTCGCGACCGCAAGTCCCAAGTATGTGTTGATGATCATCTGCGGCTCGGT	600
Db	7494	TCTACTTCAAGTCGCGACCGCAAGTCCCAAGTATGTGTTGATGATCATCTGCGGCTCGGT	7553
QY	601	AGCTCAGCTGGTTATTCTCTGCTGCTTTCTTTAAACAGATTACCGCTTCGAAAAGATTCTA	660
Db	7554	AGCTCAGCTGGTTATTCTCTGCTGCTTTCTTTAAACAGATTACCGCTTCGAAAAGATTCTA	7613
QY	661	AGAATAATACATTTCAAGCTATTTTAGACATGGCCACATATCCGTCTATGTTAGTACT	720
Db	7614	AGAATAATACATTTCAAGCTATTTTAGACATGGCCACATATCCGTCTATGTTAGTACT	7673
QY	721	TGCTAGTGTGCTAGGACTTTTTCGAAGTGTGCTGGTGGAAAAAGCTCAGTACAGAAATG	780
Db	7674	TGCTAGTGTGCTAGGACTTTTTCGAAGTGTGCTGGTGGAAAAAGCTCAGTACAGAAATG	7733
QY	781	GAAGAGTTTCAACTAGGAAAAAGCTCATACATTTTGTATAAATCATCGTTCAACGATATCA	840
Db	7734	GAAGAGTTTCAACTAGGAAAAAGCTCATACATTTTGTATAAATCATCGTTCAACGATATCA	7793
QY	841	TGGCAAGCTTGTGATTTGAAGAGTGTGGTTGATTCGAAGTTTCATCGCTTTTTCCT	900
Db	7794	TGGCAAGCTTGTGATTTGAAGAGTGTGGTTGATTCGAAGTTTCATCGCTTTTTCCT	7853
QY	901	AATGTCATAAGCACTTTTGTATACAGTGTGCTGCTGTTCTGCTGTTGCTTCTTCCTCGT	960
Db	7854	AATGTCATAAGCACTTTTGTATACAGTGTGCTGCTGTTCTGCTGTTGCTTCTTCCTCGT	7913
QY	961	GATGAGATGAGTGAATCAAGTTGTTGCAATGTTTTTGGCCATCTGGGGGATTTGTTTCT	1020
Db	7914	GATGAGATGAGTGAATCAAGTTGTTGCAATGTTTTTGGCCATCTGGGGGATTTGTTTCT	7973
QY	1021	TATGTTTATCAGCATTTATGTCATGATAGAAGCCAGAGAAGCAAGAGCTTCCTCAG	1080
Db	7974	TATGTTTATCAGCATTTATGTCATGATAGAAGCCAGAGAAGCAAGAGCTTCCTCAG	8033
QY	1081	TCTAAGAAGAGAAGAAACAAAAACAAGTAGATACCATTTGTTCCAAAGCTTAGGCAAG	1140
Db	8034	TCTAAGAAGAGAAGAAACAAAAACAAGTAGATACCATTTGTTCCAAAGCTTAGGCAAG	8093
QY	1141	ATCCA 1145	
Db	8094	ATCCA 8098	
RESULT 4	ATCHRIV48		
LOCUS	194143 bp	DNA	linear
DEFINITION	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 48.		
ACCESSION	AL161548		
VERSION	AL161548.2	GI:7268604	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE	1 (bases 1 to 111084)		

AUTHORS	Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL	Unpublished
REFERENCE	2 (bases 107966 to 194143)
AUTHORS	Hilbert,H., Braun,M., Holzer,B., Brandt,A., Duesterhoeft,A., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 194143)
AUTHORS	EU Arabidopsis sequencing, project.
TITLE	Direct Submission
JOURNAL	Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, project
AUTHORS	Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
JOURNAL	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
COMMENT	this fragment has an overlap with ATCHRIV47 at the 5' end and an overlap with ATCHRIV49 at the 3' end.
FEATURES	Location/Qualifiers
source	1..194143
	/organism="Arabidopsis thaliana"
	/molecule="genomic DNA"
	/variety="Columbia"
	/db_xref="taxon:3702"
	/chromosome="4"
	6377..9293
gene	/gene="AT4g18120"
CDS	join(6377..6796,7122..7238,7347..7501,7578..8456,8543..8627,8716..8784,8871..9293)
	/gene="AT4g18120"
	/note="similarity to Mei2-like protein, Arabidopsis thaliana, gb:D86122
	contains EST gb:W43784, AT998063.1"
	/codon_start=1
	/product="putative protein"
	/protein_id="CAB78814.1"
	/db_xref="GI:7268605"
	/db_xref="GOA:Q9SVV9"
	/translation="MASGLDHFSGIGNMLDDGDSHPIGNMLPDBEELFSGMLDDLNLSSLPATLDLEDYDLFGSGGLELETDYPDLSLNKSPRMGFADSNVDNVPQNIQFQVQSGVSSFGSLNSHPHSHVETLSGSEFLWGSPPSSAWPVNPFSNNRHNRPYSAQNSGLHLLHIGSAPGFFPRSPETSMGSAVFRGASNMNQRNLRNRETSPNFKMLSAPRRSLFTGNGSYLWPAATMVSIDDDPLEDSNQOFSNGNOADIKIQFOLDLSKIMRGEPRTLTKNIKNKYTRNMLAAIDENKSTYDLYLPIDFKNKNVGYAFINMVS PKETIAYEAFNKKWKDKFNSEKVASLAYARIQKKAALIAHFQNSSLMNEDRRRCQPIVFDGSEKYPVNNLLQQTHDLKMIYILFCVYLMNWLITCLTFETRSSEGEFAARASNSGTFLIPLSARHIKADSQNVHVIASGLKKEKYQFLG"
exon	6377..6796
	/gene="AT4g18120"
	/number=1
intron	6797..7121
	/gene="AT4g18120"
	/number=1
exon	7122..7238
	/gene="AT4g18120"
	/number=2
intron	7239..7346
	/gene="AT4g18120"
	/number=2
exon	7347..7501
	/gene="AT4g18120"
	/number=3
intron	7502..7577
	/gene="AT4g18120"
	/number=3


```
Db 35701 CTACTTACCTTCCTGTCTCAACTTCTCTTTGATCTCTGGTGCAGTTCGGCTTTTAAC 35760
Qy 421 GCCGTCCTTCTTACTTCTCTAACTCAAAAAATCACACCATTATATCAATTCATT 480
Db 35761 GCCGTCCTTCTTACTTCTTAACTCAAAAAATCACACCATTATATCAATTCATT 35820
Qy 481 GTTCCTTAAACCATATCTTCTACACTTCTTCTTATCCAACTGACCAAGATCTCCCTCT 540
Db 35821 GTTCTCTTAAACCATATCTTCTACACTTCTTCTTATCCAACTGACCAAGATCTCCCTCT 35880
Qy 541 TCTACTTCAAAAGTCGGCAAGTCCAAAGTATGTGATTGGATACATCTGCGCGGTGGT 600
Db 35881 TCTACTTCAAAAGTCGGCAAGTCCAAAGTATGTGATTGGATACATCTGCGCGGTGGT 35940
Qy 601 AGCTCAGCTGGTATCTCTGGTCTTCTTTTAAACAGATTCAGGCTTCGAAAAATCTA 660
Db 35941 AGCTCAGCTGGTATCTCTGGTCTTCTTTTAAACAGATTCAGGCTTCGAAAAATCTA 36000
Qy 661 AAGAAATACATTCAGGCTATTTTACAGATGCGCATATCCGTCTATGGTAGCTACT 720
Db 36001 AAGAAATACATTCAGGCTATTTTACAGATGCGCATATCCGTCTATGGTAGCTACT 36060
Qy 721 TGTGTAGTGTGGTAGGACTTTTGGAAAGTGGTGGTGGAAAAAGCTGAGTACAGAAATG 780
Db 36061 TGTGTAGTGTGGTAGGACTTTTGGAAAGTGGTGGTGGAAAAAGCTGAGTACAGAAATG 36120
Qy 781 GAAGAGTTTCAACTAGGAAAAAGCTATACATATTTTGATAAAACATCGGTTCAACGATATCA 840
Db 36121 GAAGAGTTTCAACTAGGAAAAAGCTATACATATTTTGATAAAACATCGGTTCAACGATATCA 36180
Qy 841 TGGCAAGCTCTTGTGATTTGGAAGTGTGGTTGATTTATCGAAGTTTCATCGCTTTTTC 900
Db 36181 TGGCAAGCTCTTGTGATTTGGAAGTGTGGTTGATTTATCGAAGTTTCATCGCTTTTTC 36240
Qy 901 AATGTCATAAGCACTCTTGTGTTTACCAGTTGTGCTGTCTTGTGTTGTCTTCTTCGGT 960
Db 36241 AATGTCATAAGCACTCTTGTGTTTACCAGTTGTGCTGTCTTGTGTTGTCTTCTTCGGT 36300
Qy 961 GATGAGTAGTGGAAATCAAGTTGGTTCGAATGTTTTTGGCCATCTGGGAAATTTGTTCT 1020
Db 36301 GATGAGTAGTGGAAATCAAGTTGGTTCGAATGTTTTTGGCCATCTGGGAAATTTGTTCT 36360
Qy 1021 TATGTTATACGATATGTCATGATAGAAAGCCAGAAAGACCAAGAGCTTCCTCAG 1080
Db 36361 TATGTTATACGATATGTCATGATAGAAAGCCAGAAAGACCAAGAGCTTCCTCAG 36420
Qy 1081 TCTAAAGAAGAAAGAAACAAAAACAAGTAGATACCATTCATGTCCAGCTTAGGCAAG 1140
Db 36421 TCTAAAGAAGAAAGAAACAAAAACAAGTAGATACCATTCATGTCCAGCTTAGGCAAG 36480
Qy 1141 ATCCA 1145
Db 36481 ATCCA 36485

RESULT 5
AP006383/c
LOCUS
DEFINITION
  Lotus corniculatus var. japonicus genomic DNA, chromosome 6,
  clone:ljT05120, TM0245, complete sequence.
ACCESSION
  AP006383
VERSION
  AP006383.1 GI:31581014
KEYWORDS
  HTG.
SOURCE
  Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
  Lotus.
REFERENCE
  1
  Asamizu,E., Kato,T., Sato,S., Nakamura,Y., Kaneko,T. and Tabata,S.
  Structural Analysis of a Lotus japonicus Genome. IV. Sequence
```

```
JOURNAL
REFERENCE
  DNA Res. (2003) In press
  2 (bases 1 to 73179)
AUTHORS
  Sato,S.
TITLE
  Direct Submission
JOURNAL
  Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute,
  Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
  Chiba 292-0818, Japan (E-mail:ssato@kazusa.or.jp,
  URL:http://www.kazusa.or.jp/, tel:81-438-52-3935(ex.2337),
  Fax:81-438-52-3934)
FEATURES
  source
    1. 73179
    /organism="Lotus corniculatus var. japonicus"
    /mol_type="genomic DNA"
    /variety="japonicus"
    /db_xref="taxon:34305"
    /chromosome="6"
    /clone_lib="ljT05120"
    /note="TAC clone:TM0245-synonym: Lotus japonicus"
  ORIGIN
    Query Match 32.9%; Score 376.6; DB 8; Length 73179;
    Best Local Similarity 61.7%; Pred. No. 8.3e-81;
    Matches 620; Conservative 0; Mismatches 379; Indels 6; Gaps 1;
    Qy 59 ACAAGTACTCTGGAGGTTAAGAGTGTCTCTATGTCACCTCTTCTTACCTGAGGAGA 118
    Db 17242 AAAGGTATTACCGTTGGTTTCGATTAGCCCTCTATGCAGCACTGGTGTGGTGC 17183
    Qy 119 CAATAGCACTCTCTTAGGTAGACTTTACTACGAAAAAGCGGTAAAGCACATGGCTCG 178
    Db 17182 CAGCTGCAATFACTTCTGGGAAGATTGTACTATGAAAAGAGGAAACAGCAATGGATG 17123
    Qy 179 AAACCTTGGTTCAGCTTGTAGGGTTTCCTTTAAACCCCTTCCTTCTCTATTTACTTAAAGC 238
    Db 17122 CAACACTAGCAACCTTGCAGGTTTCCTTATTTCTGCTTCCTTACTATCTCATCTTAGCAT 17063
    Qy 239 CTGAGCGGTCGAAGCTAAACCATTTACGAAAAAATCTCTTCTCTCTTGCACACTAT 298
    Db 17062 CCAAGGATCTCAACACCAAAATAGCTTTTCATCAGAACTCAACCTCTGTTCATGTAG 17003
    Qy 299 CTTTAGTGATATTTGGACTTGGCTTGTCTGTGGACATTTGATTTTCTTACTCATTTG 358
    Db 17002 CATTTGCTAAGTCTCCCTTGGCTTCTGTGGCAATAGATTGTTATTTACTCTCTTG 16943
    Qy 359 GGCTACTTTTACCTTCTCTCTCAACTTCTCTTTGATCTCTGGCTGCAATTTGGCTTTTA 418
    Db 16942 GACTTTTATACCTCCCTGTCTCTACTTACACACTCATTTGGCTCATCTCAATTTGGCTTTCA 16883
    Qy 419 ACGCGTCTTCTCTTACTTCTTAACTCAAAAAATCACACCATTATATCAATTCAC 478
    Db 16882 ATGCTCTCTTCTCCTTCTCAATTCACCTGAAAAATCACACCTCTCATCACTCAACTCTC 16823
    Qy 479 TTGTTCTTTAAACCATATCTTCTACACTTCTTGTATTCACACATGAACACAGAAATCTCCCT 538
    Db 16822 TAGTCTTCTAAACATTTCTCTCCCTCTCTGTTGTTTCAAAA-----GGACTCATCAT 16769
    Qy 539 CTTTCTACTTCAAAAGTCGGCAGCCCAAGTCCAAGTATGTGATTGGATACATCTGCGCGGTG 598
    Db 16768 CTGATGATTCCGGCGAAGTCTCCAAACAAAATATGATGTTGGATTATATGACACAGTTG 16709
    Qy 599 GTAGCTCAGCTGGTATTCTCTGGTGTCTTCTTTTAACAGATTACGCGTTTCAAAAGATTC 658
    Db 16708 CTGCAATCAGCAGGCTATGGATTGGTGTCTTCCCTTACCCCAAGTGTCTTCAACAGGTTA 16649
    Qy 659 TAAAGAAATACATTCGCAAGCTATTTTAGACATGGCCACATATCCCTCTATGGTAGCTA 718
    Db 16648 TTAAGAGGAACTTTCGAAAGTGGTATGATATGATTAATCTATCACTGCTCTTGTAGCTA 16589
    Qy 719 CTTGTGTAGTGTGGTAGGACTTTTGGAAAGTGGTGGGAAAGCTTCAAGCTGAGTACAGAA 778
```

Db 35701 CTACTTACCTTCCTGTCTCAACTTCTCTTTGATCTCTGGTGCAGTTCGGCTTTTAAC 35760

Qy 421 GCCGTCCTTCTTACTTCTCTAACTCAAAAAATCACACCATTATATCAATTCATT 480

Db 35761 GCCGTCCTTCTTACTTCTTAACTCAAAAAATCACACCATTATATCAATTCATT 35820

Qy 481 GTTCCTTAAACCATATCTTCTACACTTCTTCTTATCCAACTGACCAAGATCTCCCTCT 540

Db 35821 GTTCTCTTAAACCATATCTTCTACACTTCTTCTTATCCAACTGACCAAGATCTCCCTCT 35880

Qy 541 TCTACTTCAAAAGTCGGCAAGTCCAAAGTATGTGATTGGATACATCTGCGCGGTGGT 600

Db 35881 TCTACTTCAAAAGTCGGCAAGTCCAAAGTATGTGATTGGATACATCTGCGCGGTGGT 35940

Qy 601 AGCTCAGCTGGTATCTCTGGTCTTCTTTTAAACAGATTCAGGCTTCGAAAAATCTA 660

Db 35941 AGCTCAGCTGGTATCTCTGGTCTTCTTTTAAACAGATTCAGGCTTCGAAAAATCTA 36000

Qy 661 AAGAAATACATTCAGGCTATTTTACAGATGCGCATATCCGTCTATGGTAGCTACT 720

Db 36001 AAGAAATACATTCAGGCTATTTTACAGATGCGCATATCCGTCTATGGTAGCTACT 36060

Qy 721 TGTGTAGTGTGGTAGGACTTTTGGAAAGTGGTGGTGGAAAAAGCTGAGTACAGAAATG 780

Db 36061 TGTGTAGTGTGGTAGGACTTTTGGAAAGTGGTGGTGGAAAAAGCTGAGTACAGAAATG 36120

Qy 781 GAAGAGTTTCAACTAGGAAAAAGCTATACATATTTTGATAAAACATCGGTTCAACGATATCA 840

Db 36121 GAAGAGTTTCAACTAGGAAAAAGCTATACATATTTTGATAAAACATCGGTTCAACGATATCA 36180

Qy 841 TGGCAAGCTCTTGTGATTTGGAAGTGTGGTTGATTTATCGAAGTTTCATCGCTTTTTC 900

Db 36181 TGGCAAGCTCTTGTGATTTGGAAGTGTGGTTGATTTATCGAAGTTTCATCGCTTTTTC 36240

Qy 901 AATGTCATAAGCACTCTTGTGTTTACCAGTTGTGCTGTCTTGTGTTGTCTTCTTCGGT 960

Db 36241 AATGTCATAAGCACTCTTGTGTTTACCAGTTGTGCTGTCTTGTGTTGTCTTCTTCGGT 36300

Qy 961 GATGAGTAGTGGAAATCAAGTTGGTTCGAATGTTTTTGGCCATCTGGGAAATTTGTTCT 1020

Db 36301 GATGAGTAGTGGAAATCAAGTTGGTTCGAATGTTTTTGGCCATCTGGGAAATTTGTTCT 36360

Qy 1021 TATGTTATACGATATGTCATGATAGAAAGCCAGAAAGACCAAGAGCTTCCTCAG 1080

Db 36361 TATGTTATACGATATGTCATGATAGAAAGCCAGAAAGACCAAGAGCTTCCTCAG 36420

Qy 1081 TCTAAAGAAGAAAGAAACAAAAACAAGTAGATACCATTCATGTCCAGCTTAGGCAAG 1140

Db 36421 TCTAAAGAAGAAAGAAACAAAAACAAGTAGATACCATTCATGTCCAGCTTAGGCAAG 36480

Qy 1141 ATCCA 1145

Db 36481 ATCCA 36485

RESULT 5
AP006383/c
LOCUS
DEFINITION
 Lotus corniculatus var. japonicus genomic DNA, chromosome 6,
 clone:ljT05120, TM0245, complete sequence.
ACCESSION
 AP006383
VERSION
 AP006383.1 GI:31581014
KEYWORDS
 HTG.
SOURCE
 Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
 Lotus.

REFERENCE
 1
 Asamizu,E., Kato,T., Sato,S., Nakamura,Y., Kaneko,T. and Tabata,S.
 Structural Analysis of a Lotus japonicus Genome. IV. Sequence

Db 16588 CTTGTGCTTCTAGTGGGACTTTTTCGCAAGTGGAGTGGAGGGTTTGAAGAAAGAA 16529
QY 779 TGAAGAGTTTCAACTAGGAAAGCTCATACTTTTGTATGAAACATCGGTTCAACGATAT 838
Db 16528 TGGAGGGTATGAGTTGGGTAGGCTTCTATGACTCTGACTCTTTCACAGCTATAC 16469
QY 839 CATGGCAAGCTTGTGATTTGGAAGTGTGGTTGATATCGAAGTTTCATCGCTTTT 898
Db 16468 TTTGGCAAGTCCTTAGTATTTGGTAGTGGGACTGATTTATGAGGTCTCTGCACTCTTCT 16409
QY 899 CCAATGTATAAGCACTCTTTTGTATACAGTGTGCTGTCTTGTGTGTCTTCTTCC 958
Db 16408 CTATGCTATAGCACTTTGGGGGCTTAACTCTATATTTGGCTGTGATGTTCTTTC 16349
QY 959 GTCATGAGTGTAGTGAATCAAGTTGGTTCGAATGTTTGGCCATCTGGGATTTGTTT 1018
Db 16348 ATGACAGATGTTGGAATCAAGATTATCTATGTTGGTGTGCTGCTGGGATTTATAT 16289
QY 1019 CTTATGTTATCAGCATTTATGTCATGATGAGAAGCCAGAGAAG 1063
Db 16288 CTTATGTTATCAGCAGTACTCGGATGAATGCAAGTCCCAACAG 16244

RESULT 6

AF370622
LOCUS
DEFINITION Arabidopsis thaliana putative protein (T9A21.60) mRNA, complete cds.

ACCESSION AF370622
VERSION AF370622.1 GI:13877726
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1106)

REFERENCE Lam,B., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

TITLE Direct Submission
JOURNAL Submitted (17-APR-2001) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

COMMENT e-mail for correspondence: arab@sequence.stanford.edu

FEATURES This clone was isolated by RT-PCR.

source

Location/Qualifiers
1..1106
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="IV"
/clone="C531"
/ecotype="Columbia"

Gene

1..1106

/gene="T9A21.60"

657..1106

/gene="T9A21.60"

/codon_start=1

/product="putative protein"

/protein_id="AAK43941.1"

/db_xref="GI:13877727"

translation="MDMIYVSLVASCVSUWGLFASSEWKTLSSEMDNYKHGKVSIM

NLWVATVQVPSIGTGLIFELSLFSLNVLGLPVVPIILAVIFDKNGLKVIS

MILAIWFTSYVYQQIQLDDNKLKKNHEITTESPPPEAEESTWQSK"

ORIGIN

Query Match 28.0%; Score 320.6; DB 8; Length 1106;

Best Local Similarity 57.9%; Pred. No. 5.2e-67;

Matches 616; Conservative 0; Mismatches 429; Indels 18; Gaps 2;

QY 57 TCACAAAGTACTCTTGAGGTTAAGAGTGTCTCTCTATGTCACCTCTCTTCTAGCTGGAGA 116

RESULT 7

Db 47 TAACACATACAAACGCTGGCTCAGGCTGACTTTATATACATCTTTGTCAATTCAGGCCA 106
QY 117 GACAATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAGGCGTAAAGACACATGGCT 176
Db 107 AACAGTTGCTACAAATTTTGGGAGAGTATCTATGACACGGAGAAACAGTAATATGGCT 166
QY 177 CGAAACCTTTGGTTCAGCTTTGTAGGGTTTCTTTTAAACCTTCTCTGCTATTTATTTACTTAAA 236
Db 167 AGCAACGGTAGTTCAACTTTTGGCTTCTCTGCTACTTTCCATATATATCTTTGTCATT 226
QY 237 GCTGAGCGCTCAAGACATAAAACCAATACGAAAAAACTACTTCTTCTCTTGTACACT 296
Db 227 TAAAAACACATGCAACAACATAGAGATGGAAGAAAGAACCTCACCT-----AGGAACCG 280
QY 297 ATCTTTAGTGTATATTTGGACTTTGGCTTGTCTGCTGCGACATTTGATTTTGTACTCATTT 356
Db 281 TGTATTTGGTTTACGTAGTGTCTTGGACTTCTTGTAGGACAGATTTGCTATCTGTACTCCAT 340
QY 357 TGGGCTACTTTAACCTTCTCTCAACTTTTCTCTTTGATCTCTGGTCGCAATTTGGCTTT 416
Db 341 TGGACTTCTTTACTTACCCTTTTCTACCTATTTCCCTGATCTCTGCACTCTCAGTTAGCCTT 400
QY 417 TAAACGCGCTTCTCTTACTTCTTAACTCAACAAAAATCAACANATTTATCTCAATTC 476
Db 401 CAATGCTTTCTCTCTATTTTCTTAACTCACAACAACTTACCCCTATCATTTTAAATTC 460
QY 477 ACTTGTCTCTTAAACCATATCTTACACTTCTGTTATCCAACTTCAACAGTGAACCAAGATCTCC 536
Db 461 TCTTTTCTCTTAACTATATCTTCCACCTTCTTCCCTTCAATTAATGAGGAGACAGACTC 520
QY 537 CTCTCTTACTTCAAAAGTCCGACGCAAGTCCAAAGTATGATTTGGATACATCTGGCGGT 596
Db 521 CACAAAAGTTACAAAAGGAG-----AGTATGTCAAAAGTTTCTATATGACCCGT 568
QY 597 CGGTAGCTCAGCTGGTGTATTTCTCTGGTGTCTTTTAAACAGATTTAGCGCTTCGAAAGAT 656
Db 569 TGCTCGCTTCTGCTGGTTATGTTCTATCTTATCCCTACAAACAGTAGCCCTTTCTTAAAGT 628
QY 657 TCTAAGAAATACACATTTCAAGGCTATTTAGACATGGCCACATATCCGCTCTATGTTAGC 716
Db 629 CCTAAGAAAGCAAAATTTCTCAGAAGTTATGATATGATAATCTACGTAGTCTAGTGGC 688
QY 717 TACTTGTGTAGTTGTGGTAGGACTTTTGGAGTGGTGGGTGGAAAAAGCTGAGTACAGA 776
Db 689 CAGTTGTGTAGCGTGGTGGGCTTTTGTAGCAGTGAAGTGAAGAACTTTGAGCAGTGA 748
QY 777 AATGAAGAGTTTCAACTAGGAAAGCTCATACTATTTTGTATGATAAATCGGTTCAACGAT 836
Db 749 AATGGATAACTACAAACATGGGAAGGTATCTTACATTTATGAACTTAGTGTGGACAGCTGT 808
QY 837 ATCATGGCAAGCTTCTTTGATTGGAAGTGTGGTTTGTATTCGAAGTTTTCATCGCTTTT 896
Db 809 TACCTGGCAGGATTTCTCATCGTGGGACAGGACTGATCTTCAGCTCTCTCTCTATT 868
QY 897 TTCCAATGTCTAAGCACTCTTTTGTATACAGTGTGCTGTGCTGTCTTGTCTTCTTCT 956
Db 869 CTCAAATGCAATTAAGGTTTGGGACTCCAGTGGTTCCTATCTTGGCTGTAAATCATTTT 928
QY 957 CGGTATGAGATGAGTGGAAATCAAGTTGGTTGCAATGTTTTTGGCCATCTGGGGATTTGT 1016
Db 929 CCATGACAAAATGAATGGGTTAAAGGTGATTTCTATGATCTCTAGCTATTTTGGGGTTTCA 988
QY 1017 TTCTTATGTTATCAGCATTTATGTCATGATAGAGGCGACAGAGAGACCAAGAGCTTCC 1076
Db 989 TTCTATGTCTACCAACAATATCTTGTATGACAAAACCTTGAAGAAAAATCATGAATCAC 1048
QY 1077 TCAGTCTTAAAGAAAGAAAGAAACAAAAACAAGTAGATACCAATT 1119
Db 1049 AACACAGATATCCCTTGACCCACACAGAGCAGAGAGTCAACT 1091

AY096558
LOCUS AY096558 1165 bp mRNA linear PLN 18-SEP-2002
DEFINITION Arabidopsis thaliana putative protein (At4g18200) mRNA, complete cds.
ACCESSION AY096558
VERSION AY096558.1 GI:20465496
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1165)
AUTHORS Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Open Reading Frame (ORF) Clones
TITLE Arabidopsis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1165)
AUTHORS Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
TITLE Arabidopsis
JOURNAL Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.
Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.
FEATURES
Location/Qualifiers
1..1165
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U18628"
/ecotype="Columbia"
/note="This clone is in pUNI 51."
1..1165
/gene="At4g18200"
1..1134
/gene="At4g18200"
/codon_start=1
/evidence=experimental
/product="putative protein"
/protein_id="AAM20208.1"
gene
CDS
1..1165
/gene="At4g18200"
ORIGIN
3'UTR
Query Match 27.8%; Score 318.4; DB 8; Length 1165;
Best Local Similarity 57.8%; Pred. No. 1.8e-66;
Matches 593; Conservative 0; Mismatches 421; Indels 12; Gaps 1;
Qy 28 GAAGGGAAATTTTCAACAGAGAGAGAGAGTCAAGTACTCTTGGAGGTTTAAAGAGTGTCT 87
Db 70 GAAACTGAATCAATTTTCAGTACTCTCAACGAGAACTGTAAGAGGTGGCTCCGTTCTCC 129
Qy 88 CTCTATGTCACTCTCTCTTAGCTGGAGAGACATAGCCACTCTCTTAGGTAGTACTTTTAC 147
Db 130 ATATACGCAATCTTTGTCACTCTTGCACCACTCTGTACAGTCTCTGGGTAGACTGTAC 189
Qy 148 TACGAAAAGCGGTAAAGCACATGCTCGAAACCTTGGTTCAGCTTGTAGGGTTTCCT 207
Db 190 TATGAAATGGAGGAAAGACACATATGTGTACACTCTTCTCAACTCAATTTGGGTTCCCT 249
Qy 208 TTAACCCCTTCTTCTGCTATTATTACTTAAAGCCTGAGCGCTCAAGACTTAAACCAATTACC 267
Db 250 GTACTGATTTCTTCCGCTCTCTTCTCGAATCAGGCAACCCAAATCAACAGATACAAAT 309
Qy 268 AAAAAAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 327
Db 310 TTCAGTCAGTCCCTCTCTTCCACCACTTGCATCGGTTTACTTGTGCATCGACTGCTA 369
Qy 328 GTTCTCGACATTTATTTTGTACTCAATTTGGGTACTTTTACCTTCTGTCTCAACTTTC 387
Db 370 GTGTCTCTTATGCTTATTTGTCTGAGTTGGGTGCTTCTTCTTCTTCTTCTTCTTCTTCT 429
Qy 388 TCTTTGATCTCTGCGTCGCAATTTGGCTTTTAAAGCCGCTTCTTCTTCTTCTTCTTCTTCT 447
Db 430 TCCCTCATCTTGGCTCACAGTTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 489
Qy 448 CAAAAATCACACCAATTTATCTCAATTTCACTTGTCTCTTCAACCATATCTTCTTCTTCT 507
Db 490 CAAAAGTTCACTCTTCTTGTAGTCAATTTCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 549
Qy 508 CTTGTTATCCAACATGAACAGAACTCTCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 567
Db 550 CTCTGCTCAACTGATTCAGAA-----NACACACAAATGATCTTAGAGTA 597
Qy 568 AAGTATGTGATGGATACATCTCGGGGTGCGTAGCTCAGCTGGTTATCTCTGGTGCTT 627
Db 598 CAATATGTGATGGGTTCACTCTGTACCATTTGGTGTCTCCGCTGGGATTTGGATTGGTACTA 657
Qy 628 TCTTTACAGATTACGGGTTTGGAAAGATTTCTAAGAAATACACATTTCAAGGCTATTTTA 687
Db 658 TCTCTGATACAACTGCTCTTTTTCAGGAAAGTTTTCAGGAAAGATATACATCTCAGCAGTCTG 717
Qy 688 GACATGGCCACATATCCGTTCTAGTGTAGTCTTGTGTAGTGTGGTGGAGCTTTTGGGA 747
Db 718 GACTTGGCCAAATTACCAGTCTCTAGTTGCAACTTTGTGTGCTACTCATAGGACTGTTTGA 777
Qy 748 AGTGGTGGGTGGAAAAAGCTGAGTACAGAAATGGAAGAGTTTCAACTAGGGAAGGCTCA 807
Db 778 AGTGGAGAGTGGAGAACTCTGCCAAAGTGAGATGAGAAATACAAACTGGGAAAGTGTCA 837
Qy 808 TACATTTTGTAAACATCGGTTTCAACGATATCATGGCAAGCTTGTGTGATGGAGTGT 867
Db 838 TATATCTTGTAGCTTTGGGCTCAGCAGCTATTTTCTGGCAAGTATACACTGTGTGTTGTG 897
Qy 868 GGTTTGTATTACGAAGTTTTCATCGCTTTTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 927

/db_xref="GI:20465497"
/translation="MGISOVHYCNGDQNLNLEILDHETBSFSVPQKNCKRWLRVSI
YALPFCQPLATVGLRYLGLYGLYAVAYLGLYLPVLSLILPFRFRIRQPKSTDT
NFSQSPFTTLASVYLCTGLVSAVGLLYLPVLSLILPFRFRIRQPKSTDT
LNSQKFTPLIVNSLFLTLVSSALLVNTSDENTNSRVQVYVIGFTICIGASAGILG
LSLIQLLPRKVPFTKHTSSAVLDLANVOSLVATCVLILGLFASGEMWRTLPSEMRNYKLG
KVSYLTLASAAIPWQVYTVGCGLIPESSSVFSNSITAVGLPIPVVAVIVPHDKMD
ASKIFSIIILAIWGLFSFYOHYLDLDEKLKTKCTQPKPVEETQL"
1135..1165
/gene="At4g18200"


```
QY 568 AAGTATGTGATGATACATCTCGCGGTTCGGTAGCTCAGCTGGTTATTCTCTGGTGCTT 627
Db 669 CAATATGTGATGTGGTTCACTGTATACATTTGGTGTCTCGCTGGGATGGATTTGTTCTA 728
QY 628 TCTTTAAACAGATTACCGGTTCGAAAGAAATCTTAAAGAAATACACATTCAAGGCTATTTTAA 687
Db 729 TCTCTGATACAACTGCTCTTCAGGAAAGTTTTCAGGAAGCATACATCTCAGCAGTCTCTG 788
QY 688 GACATGGCCACATATCCGCTATAGGTAGTACTTGTGTAGTTGTGGTAGGACTTTTGTGA 747
Db 789 GACTTGGCCAAATTACAGTCTCTAGTTGCAACTTGTGTGGTACTCATAGGACTGTTTGA 848
QY 748 AGTGTGGGTGGAAGAGCTAGTACAGAAATGGAAGATTTCAACTAGGGAAGCTCA 807
Db 849 AGTGAGAGTGGAGAACTCTGCCAAGTGAGATGAGAACTACAACTGGGGAAGTGCA 908
QY 808 TACATTTTGAATAACATCGGTTCAACGATATCATGCAAGCTTGTGTTGATTGGAAGTGT 867
Db 909 TATATCTTTGACTTTGGCCTCAGCAGCTATTTCTGGCAAGTATACACATGTTGGTTGTG 968
QY 868 GGTTTGATATCGAAGTTTCACTCGCTTTTTCGAATGTCATAAGCACHCTTTGTTTACCA 927
Db 969 GGAATTAATCTTCGAGTCTTCTCTGTGTCTCCAAATCCAAATCCAAATCCAAATCCAA 1028
QY 928 GTTGTGCTGTTCTTGTGCTGTCTCTTCGCTGATGATGAGTGGGAATCAAGTTGGTT 987
Db 1029 ATAGTTCAGTTGTAGCAGTATAGTTTTCATGATAGATGATGATGATGATGATGATGAT 1088
QY 988 GCAATGTTTTTGGCCATCTGGGGAATTTGTTTCTTATGTTTATCAGCAATATGTCAATGAT 1047
Db 1089 TCCATCAATTTAGCTATCTGGGCTTCTCTCAATTTGTTCTATCAGCACTACCTCGAGAA 1148
QY 1048 AGAAG 1053
Db 1149 AAGAG 1154
```

```
RESULT 9
AX507517 AX507517 3387 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 2212 from Patent WO0216655.
DEFINITION AX507517
ACCESSION AX507517
VERSION AX507517.1 GI:23388754
KEYWORDS
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Harper,J.F., Krops,J., Wang,X. and Zhu,T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 2212 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source
1. .3387
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 27.8%; Score 318.4; DB 6; Length 3387;
Best Local Similarity 57.8%; Pred. No. 1.6e-66;
Matches 593; Conservative 0; Mismatches 421; Indels 12; Gaps 1;
QY 28 GAAGGGAATTTTCAACAGAGAGAGAGTACCAAGTACTCTTGGAGGTTAAGAGTGTCT 87
Db 2323 GAAACTGAATCAATTTTCAGTACCTCAACGAGAACTGTAAAGAGTGGCTCCGTGTCTCC 2382
QY 88 CTCTATGTCACTCTCTCTTAGCTGGAGAGACAATAGCCACTCTCTTAGGTAGACTTTAC 147
```

```
Db 2383 ATATACGAATCTTTGTGATCTTCTGCAACACCTTGTCTACAGTTCTGGGTAGACTGTAC 2442
QY 148 TACGAAAGCGGTAAAGACATGGCTCGAAACCTTTGGTTTCAAGCTTTGTAGGGTTTCCCT 207
Db 2443 TATGAAATGAGGGAAGACATATCTGTTAAACATCTTCTTCAACTCATTTGGCTTCCCT 2502
QY 208 TTAACCCCTTCTGCTATATTACTTAAAGCCTGAGCGGTCCAAGACTAAACCAATACC 267
Db 2503 GTACTGATTTCTTCCGCTTCTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAAT 2562
QY 268 AAAAAAATCACTCTTCTTCTTCTGACACTATCTTTAGTGTATATTGGACTTTGGCTTGCTT 327
Db 2563 TTCACTCAGTCCCCTTCTTCTCACCACTGTCGTTTACTTGTGCATGGAGTGTCTA 2622
QY 328 GTTCTGAGCAATGTATTTTGTACTCATTTGGGCTACTTTACCTTCTCTGTCTCAACTTTC 387
Db 2623 GTGTCTGCTTATGCTTATTTGTCTGAGTTGGGTGCTTTACTTACCAGTCTCTACTTTTC 2682
QY 388 TCTTTGATCTCTGGTCCGAATTTGGCTTTTAAAGCGCTTCTTCTTACTTCTTAACTCA 447
Db 2683 TCCCTCATCTTGGCTCACAGTTGGCTTCACTGCTCTTTTCTCATATTTCTTAACTCG 2742
QY 448 CAAAAAATCACACCAATTTATATCAATTTCACTTCTTCTTAAACCATATCTTCTACACTT 507
Db 2743 CAAAAGTTCACTCTTGTATGATCAATTTCTTTGTTTCTCTTACGGTTTCTCTGCCCTC 2802
QY 508 CTTGTTATCCAAATGAACAGCAATCTCCCTTCTTACTTCTTCTTCTTCTTCTTCTTCTT 567
Db 2803 CTCTGCTCAACACTGATTTCAGAA-----AACACAACAAATGATCTAGAGTA 2850
QY 568 AAGTATGTGATTTGATACATCTGCGCGGTTCGGTAGCTCAGCTGGTTATTCTCTGGTGT 627
Db 2851 CAATATGTGATTTGGTTTCACTCTGTACCAATTTGGTGTCTTCCGCTGGGATTGGAATTGGTACTA 2910
QY 628 TCTTTAACAGATTACCGGTTTCGAAAGAAATCTTAAAGAAATACACATTTCAAGGCTATTTTA 687
Db 2911 TCTCTGATACAACTGCTCTTTCAGGAAGTTTTCAGGAAGCATATCTCTCAGCAGTCTGT 2970
QY 688 GACATGCCACATATCCGTTCTATGGTAGTACTTGTGTAGTGTGTGGTAGGACTTTTGGGA 747
Db 2971 GACTTGGCCAAATACCAGTCTCTAGTTGCAACTTGTGTGTACTCTATAGGACTCTTTTGA 3030
QY 748 AGTGTGGGTGGAAGAAAGCTGAGTACAGAAATGGAAGTTTCAACTAGGGAAGCTCA 807
Db 3031 AGTGAGAGTGGGAACCTCTGCCAAGTGAGATGAGAAACTACAAACTGGGGAAAGTGTCA 3090
QY 808 TACATTTTGTATAACATCGGTTCAACGATATCATGCGCAAGCTTGTGTTGGAAGTGT 867
Db 3091 TATATCTTGAATTTGGCTCAGCAGCTATTTTCTGGCAAGTATACACTGTTGGTTGTGTG 3150
QY 868 GGTTTGATTTACGAAGTTTCAATCGCTTTTTCGAATGTCTATAAGCACTCTTTGTTTACCA 927
Db 3151 GGATTAATCTTCGAGTCTTCTTCTGTGTTCTTCCAAATCCATAACTGCTGTGGGACTGCT 3210
QY 928 GTTGTGCTCTTCTTGTGCTGTCTTCTTCCGCTGAGTGTGAGTGAATCAAGTGTGT 987
Db 3211 ATAGTTCAGTTGTAGCAGTGTAGTTTTCATGATAAGATGGATGATCAAGATTTTTC 3270
QY 988 GCAATGTTTTTGGCCATCTCGGGAATTTGTTTCTTATGTTTATCAGCAATTTATGTCAATGAT 1047
Db 3271 TCCATCAATTTAGCTATCTGGGCTCTCTTTCATTTTGTCTATCAGCACTACCTCGAGAA 3330
QY 1048 AGAAG 1053
Db 3331 AAGAG 3336
RESULT 10
AX652015
LOCUS Sequence 908 from Patent WO0300098.
DEFINITION AX652015
ACCESSION AX652015
```

VERSION AX652015.1 GI:29154833
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1
REFERENCE Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
AUTHORS Katagiri, F., Qian, S., Qian, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 0300898-A 908 03-JAN-2003;
SYNGENTA PARTICIPATIONS AG (CH)
FEATURES Location/Qualifiers
source
1..3387
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 27.8%; Score 318.4; DB 6; Length 3387;
Best Local Similarity 57.8%; Pred. No. 1.6e-66;
Matches 593; Conservative 0; Mismatches 421; Indels 12; Gaps 1;
QY 28 GAAGGGAATTTTCAACAGAGAGAGAGTCAAGTCACTCTTGGAGGTTAAGAGTCT 87
DB 2323 GAAACTGAATCAATTTTCAGTACCTCAACAGAGAGAGTCAAGTCACTCTTGGAGGTTAAGAGTCT 2382
QY 88 CTCATGTGCTACTCTCTCTAGTGTGAGAGACATAGCCACTCTCTTAGGTAGACTTTAC 147
DB 2383 ATATAGCGAATCTTTGTGCTATCTTCTGCCAACCACTTGTACAGTCTCTGGGTAGACTGTAC 2442
QY 148 TAGCAAAAAGGGGTTAAAGCACATGGCTCGAAACCTTGGTTTTCAGTCTTGTAGGTTTCT 207
DB 2443 TATGAAATGGAGGGAAGACATATGTGTGAACACTTCTTCAACTCACTAGTCTTCTCT 2502
QY 208 TTAACCTCTCTGCTTATTTACTTAAAGCCGTGAGCCGTGCAAGACTAAACCAATTACC 267
DB 2503 GTACTGATTTCTCTCCGCTTCTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAT 2562
QY 268 AAAAAAATCTCTCTCTCTGACACTATCTTTAGTGTATATTTGGACTTGGCTTGGCTT 327
DB 2563 TTCAGTCACTGCTCTCTTCCACCCCTTGATCGGTTTACTTGTGCACTGAGCTGCTA 2622
QY 328 GTTGTGCAATGTATTTGTACTCATTTGGGCTACTTTTACTCTCTGCTCAACTTTC 387
DB 2623 GTGCTGCTATGCTTATTTGTCTGAGTTGGGTTGCTTACTTACCAGTCTCTACTTTC 2682
QY 388 TCTTTGATCTCTGCGTGGCAATTTGGCTTTTAAAGCCGTCTTCTTACTTCTCTAAACTCA 447
DB 2683 TCCCTCATCTTGGCCTTCAAGTTGGCCTTCACTGCTCTTTTCTCATATTTCTTAACTCG 2742
QY 448 CAAAAATCACACCAATTTATCTCAATCTTCTTCTTCTTAACTATCTTCTCACTT 507
DB 2743 CAAAGTTCACTCTTTGATGATCAATCTTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2802
QY 508 CTGTGTTATCCAAATGAGAGAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 567
DB 2803 CTGCGTCAACACTGATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2850
QY 568 AAGAT 627
DB 2851 CAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2910
QY 628 TCTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
DB 2911 TCTTGTATCAACTGCTCTTCAAG 2970
QY 688 GACATGGCACAATATCCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 747
DB 2971 GACTTGGCCAAATACAGTCTCTAGTGTGCACTTGTGTGCTTGTGTGCTTGTGTGCTTGTG 3030
QY 748 AGTGGTGGTGGGAAAAAGCTGATACAGAAATGGAAGAGTGTTCACACTAGGGAAGAGCTCA 807

DB 3031 ACTGGAGAGTGGAGAACTCTGCCAAGTGAAGATGAGAACTACAAACTGGGGAAGTGTCA 3090
QY 808 TACATTTTGTATAAATCATCGTTTCAACGATATCATGCAAGCTTGTGTGATTCGAAGTGT 867
DB 3091 TATATCTTGAATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 3150
QY 868 GGTGTTGATATCAAGAGTCTTTCATCGCTTTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 927
DB 3151 GGATTTAAATCTTGGAGTCTTCTTCTGTTTCTTCCATTCCTTCCATTCCTTGGAGTCT 3210
QY 928 GTTGTGCTGCTTCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 987
DB 3211 ATAGTTCCAGTTGTAGCAGTGTATGTTTTCATGATGAAGATGATGATGATGATGATGATGAT 3270
QY 988 GCAATGTTTTCGCGCATCTGGGATTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1047
DB 3271 TCCATCATTTTAGCTATCTGGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3330
QY 1048 AGAAAG 1053
DB 3331 AAGAAG 3336
RESULT 11
AX033548
LOCUS AX033548
DEFINITION Sequence 5 from Patent WO049152.
ACCESSION AX033548
VERSION AX033548.1 GI:10280292
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1
REFERENCE Andre, B., Buerkle, L., Frommer, W.B. and Gillissen, B.
AUTHORS Nucleic acids that code for a nucleobase transporter
TITLE Patent: WO 0049152-A 5 24-AUG-2000;
JOURNAL ANDRE BRUNO (BE); BUERKLE LUKAS (DE); FROMMER WOLFF B (DE);
GILLISSEN BERND (DE)
FEATURES Location/Qualifiers
source
1..1194
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 27.7%; Score 317.2; DB 6; Length 1194;
Best Local Similarity 57.5%; Pred. No. 3.4e-66;
Matches 596; Conservative 0; Mismatches 428; Indels 12; Gaps 1;
QY 18 GGTAGAACCTGAAGGAAATTTTCAACAGAGAGAGAGAGTCAAGTCACTCTTGGAGTGT 77
DB 75 GGTGTAACCTGAATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 134
QY 78 AAGAGTGTCTCTATGTCACTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 137
DB 135 TCGTGTCTCCATATACGTAATCTTTTGTCTCTTTTGTCCAGCCACTAGCTACAAATCTCTGG 194
QY 138 TAGACTTTTACTACGAAAAAGCGGTAAAGACACATGGCTCGAAAACTTGGTTTCACTTGT 197
DB 195 TAGATTGTACTATGAATGGAGGAAATAGACATATGTGGTAAACACTTCTTCACTCAT 254
QY 198 AGGGTTTCTCTTAAACCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 257
DB 255 TGGCTTCCCTGTACTGTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 314
QY 258 AACCATTTACAAAAAATCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 317
DB 315 AGATACAAATTTCACTGCTAGTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 374


```
Qy 1030 CAGCATTATGCAATGATAGAAAGCCAGAAAGACCAAGAGCTTCCTCA 1079
|||||
Db 1063 CAGAATCATATGATGACTTGAAGTAAAGCAAGCAGCAACCAAGCTCA 1112

RESULT 14
AY062782
LOCUS Arabidopsis thaliana Unknown protein (Atlg44750; T12C22.2) mRNA,
complete cds.
ACCESSION AY062782
VERSION AY062782.1 GI:17065411
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1372)
Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (14-NOV-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arabsequence.stanford.edu

RIVEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIVEN
Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Nguyen,M.,
Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,
Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,
Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Southwick,A., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.

FEATURES
source
1..1372
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="I"
/clone="RAFL06-11-F13"
/ecotype="Columbia"
/note="This clone is in pBluescript"
1..1372
/gene="Atlg44750; T12C22.2"
50..1189
/gene="Atlg44750; T12C22.2"
/codon_start=1
/product="Unknown protein"
/protein_id="AAL32860.1"
/db_xref="GI:17065412"
/translaton="MAKPEPRLVLTGAAGNQEPILVKESVVGIPTPLLKLKSWQWV
LVSVNIPFLIGQAAVLGFRFYDEGNSKMATLVQTAAPFPIYPIPLLLPSSAV
ESSSSCSKLVLYLVLLGVI IAGDMLVSVGLLYLSASFSYLI CATOLAFNAFVS
FINAQKFTALLINSVLLSFAALIALNDADTPSGVRSKYIVGFVCTLAASLYSL
LLSLMQFSFKILKRETFVSVLEMQIYTSILVATCVSIGDFASGEWRTLHGEWGYHK
GQASVLTFLVWATVWQCSVGVGLTFLVTSFSNVISTLSLATVTLAALVWPRDKM
SGVKIMAMLIAIWGFASVYVQNHIDDLKVRQARQAQGRVPEPC"

gene
CDS
ORIGIN
```


RESULT 15	
TI2C22	
LOCUS	115421 bp DNA linear PLN 22-JUN-2000
DEFINITION	Sequence of BAC TI2C22 from Arabidopsis thaliana chromosome 1,
ACCESSION	AC020576
VERSION	AC020576.2 GI:6728952
KEYWORDS	HTG.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 (bases 1 to 115421)
AUTHORS	Liu,S., Vayenberg,M., Sakano,H., Lee,J., Lenz,C., Pham,P., Toriumi,M., Yu,G., Chin,C., Chlou,J., Choi,E., Chung,M., Gonzalez,A., Howng,B., Liu,A., Altafi,H., Brooks,S., Buehler,E., Chao,Q., Conn,L., Conway,A., Hansen,N., Johnson-Hopson,C., Khan,S., Kim,C., Lam,B., Miranda,M., Nguyen,M., Palm,C., Shinn,P., Southwick,A., Davis,R., Ecker,J., Federspiel,N. and Theologis,A.
	The sequence of BAC TI2C22 from Arabidopsis thaliana chromosome 1
	Unpublished
JOURNAL	2 (bases 1 to 115421)
REFERENCE	Theologis,A.
AUTHORS	Direct Submission
TITLE	Submitted (05-JAN-2000) Plant Gene Expression Center, 800 Buchanan
JOURNAL	Street, Albany, CA 94710, USA
REFERENCE	3 (bases 1 to 115421)
AUTHORS	Theologis,A.
TITLE	Direct Submission
JOURNAL	Submitted (09-FEB-2000) Plant Gene Expression Center, 800 Buchanan
REFERENCE	Street, Albany, CA 94710, USA
AUTHORS	Theologis,A.
TITLE	Direct Submission
JOURNAL	Submitted (22-JUN-2000) Plant Gene Expression Center, 800 Buchanan
REFERENCE	St., Albany, CA 94710, USA
TITLE	On Jan 21, 2000 this sequence version replaced gi:6671920.
COMMENT	The sequence of BAC TI2C22 from Arabidopsis thaliana chromosome 1.
FEATURES	Location/Qualifiers
Source	1..115421
	/organism="Arabidopsis thaliana"
	/mol_type="genomic DNA"
	/cultivar="Columbia"
	/db_xref="taxon:3702"
	/chromosomes="1"
	/clone="TI2C22"
	185..1855
gene	/gene="TI2C22.1"
CDS	join(185..360,885..996,1263..1370,1697..1855)
	/gene="TI2C22.1"
	/codon_start=1
	/evidence=not experimental
	/product="TI2C22.1"
	/protein_id="AAF78257.1"
	/db_xref="GI:8655984"
	/translation="WTEGEINVLGIKEEGEINVLSTKEEGAKVFFGKLGRRGPSRS
	CNIGALIGHLDIRKFPVLVFTVITGEVDSFKLHNSIKFSCGSKWRRRYGVKKEA
	KTRDRGVEDATRKRGROVSTPANGSASVLKIYGEKLMKTSWNFTVNPDSVSKGV
	WRNLTPPEGSEDIVYAAATELYAM"
	5388..7868
gene	/gene="TI2C22.2"
CDS	join(5388..5430,6772..7868)
	/gene="TI2C22.2"
	/note="Contains similarity to purine permease from
	Arabidopsis thaliana gb AF078531. EST gb A1997301 comes
	from this gene."
	/codon_start=1
	/evidence=not experimental
	/protein_id="AAF78258.1"
	/db_xref="GI:8655985"
	/translation="WAKGFVRVLVTGAAGNOBILVKEESVVGIFPTPLLKLSQWQWV
	LVSNNIFLIQGQAASVLLGRFYDEGGNKMATLVQTAAFPILYIPLLLPSASV
	ESSESSCLKYIVLIYVLLGVIIAGDNLMSVGLLYLSASTYSLICATQLAFNAVFSY
	FINAQKFTALINSVLLSFSAAALITSLDADTDSGVSRSKYIVGFCVTLTAASALYSL
	LLSLMQFSEFKILKRETFVSELEMOIYTSLVATCVSIVGLPASGEMRTLHGMEGYHK
	GOASVLLTAVMTWQVCSGVVGLIELVTSLESNVISTLSLATVPLAALVVPDRKM
	SGVKIMAMLIAGVFPASVYQNHIDDLKVRQARQAQGRVEPPC"
	complement(9594..11127)
gene	/gene="TI2C22.3"
CDS	complement(join(9594..9744,9836..9957,10759..11127))
	/gene="TI2C22.3"
	/note="Contains similarity to hypothetical protein
	F19B11.17 gi 4406763 from Arabidopsis thaliana BAC F19B11
	gb AC006836."
	/codon_start=1
	/evidence=not experimental
	/protein_id="AAF78259.1"
	/db_xref="GI:8655986"
	/translation="MQFSRNSILRQLSRKEGWSRASKRWTSKSDSTAFNDTSGGGYS
	SMEGYGVISGDTAARSKRVVVDESSRSKHAMWALTHTLTNGDILVTLHHVVSPL
	DEATPSLAQSLGSLCKACKPEVDVIQGPKLATVLSQVKLEVSIVLVLGQKKSAP
	LISLCGFSRSEELVNRINCINGADCLITIGVRKQCKGVGGYLINRQKNFWLLA"
	complement(20534..22109)
gene	/gene="TI2C22.4"
CDS	complement(join(20534..20615,20695..20803,21229..21276,
	21381..21566,21643..21760,21843..22109))
	/gene="TI2C22.4"
	/note="Contains weak similarity to tail completion
	gi 5354213 from coliphage T4 gb AF158101. ESTs
	gb AA650799, gb AA041054, gb R29873, gb AA712908 come from
	this gene."
	/codon_start=1
	/evidence=not experimental
	/protein_id="AAF78260.1"
	/db_xref="GI:8655987"
	/translation="WTSHAJEVTKVLEADVAVTAVETVHHHHHHQDENHSTNPI
	SDPRDREALRQENRRRLTLESNLKLFETLAESAASFSDCPSDLARLVMTVTSRD
	FLAKLENRQALSNGTQPFKPTEDDVKTVEVLIMDHQEPSSMWLVTDMMVPSN
	VEQSAIDNEHYIVNBEHVDAVAHLKATGFKLQNLKPEELQKLVQEVYALSKV
	GKVVDIWHAGKWFYTLSTWLAFGVLQARGVLKIAKGVHATSKVLRAL"
	22453..24681
gene	/gene="TI2C22.5"
CDS	join(22453..22580,22790..22895,22987..23237,23328..23433,
	23515..23832,23910..23989,24078..24202,24407..24681)
	/gene="TI2C22.5"
	/note="Contains similarity to hypothetical protein
	TI2GL3.150 from Arabidopsis thaliana gi 7486671."
	/codon_start=1
	/evidence=not experimental
	/protein_id="AAF78261.1"
	/db_xref="GI:8655988"
	/translation="METAEMNHNGSKNLKNVDATEIEFKILAAALRSRVTLRNEADC
	FTLVSRMLEEDIGLEKCDLVYKSFVKEHLVKCLEAGNNDTSENSQETEREDEI
	PTKEVAQSEHEHPMDAGENTSREAKDVKGKNETLQORDIKALRKASYYIKAN
	SETITMASRLLEEDLKEKESLDLFAFKINKELDDEVLPDPAKPKSTESIVKNVKK
	KVSTPSKMSVSEYNSDSTEGNVNEEVAVKTMARKVLKSPKEMMGKRSKNGKV
	SGRRKAKHEIDSDSDSGSEKSLQKTETATDVYGRVHLKSLKSCGMSVPPN
	IYKKAQAPOEKREAMLIELEQILAKEGLSSDPSALRIKEVKRKNISRLGKIDTN
	NIVNRSRRSSTSPAPPKPVTAESSESEDEPESNEEESNEKAERGSGSSEVEE
	SNSCENLYLTSK"
	complement(25103..26276)
gene	/gene="TI2C22.6"
CDS	complement(join(25103..25212,25310..25389,25470..25541,
	25625..25668,25739..25795,26013..26089,26174..26276))
	/gene="TI2C22.6"
	/note="Contains similarity to cation transport protein
	CHAC from Escherichia coli gi 2506988. ESTs gb AA605474,
	gb A1995104, gb R90162 come from this gene."
	/codon_start=1
	/evidence=not experimental
	/protein_id="AAF78262.1"
	/db_xref="GI:8655989"
	/translation="MMVFGYGLIWKTFPPFDESLPGFIKGYRRRVHOGSTHRRGTPD
	FPGRVTLAEAAHEEVKLEVRKQYDQKEYLDLFTQDSNASEPAGVMVYIAGSPDK
	SNNYLGPAFLPIAKQITVAKGPSGNRDYLFNLEEALAQLFKDKHVTDLANQVRH

gene	ILSSEELDIDATAATANNV complement(27042..29558) /gene="T12C22.7"	Qy	250	AAGACTAAACCATTACCAAAAAAATACTCTTCTCTTGACACTATCTTTAGTGTAT	309
CDS	complement(join(27042..27234,27331..27482,27571..27729, 27804..28044,28296..28412,29221..29283,29371..29558)) /gene="T12C22.7"	Db	7023	TCGGTAGTGTAGAGTCTTCGGAGAGTTCGTGTTCACTCAAGTACATGTTTGTGATCTAT	7082
	/note="Contains similarity to Mtn21 gene from Medicago truncatula gb Y15293 and contains two integral membrane protein domains DUF6 of unknown function PF 00892. ESTs gb A1998702, gb 230851 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAF78263.1" /db_xref="GI:8655990"	Qy	310	ATTGGACTTGGCTTGGTGTGGCAATGTATTTGTACTCATTTGGGCTACTTTTAC	369
	/translation="MKGSMKIRPILAILSLQFYAGMVIITWVSPKMGMDHWLVT YRHVATVWAPFALMERKIPKOTLAIWRLLALGILPELMDONLIYIIGLKNVTAS YTSATFALPAPFILIARLETNFRKVSIAKVVTGVTITVGGAMIMLYIKGPALIE IVKAHNSFHGSSSTPGQHVLTGTAIMGSIWAAFFLIQSYTLKVPYPAELSLVT LICGIGTILAIASLIMVRDPSAWKIGMDSGLAAVSGVCSGIAVYIOSIVIKORG PVFTTSPCMIIITAFGLALVLAEKIHLASILGAVFIVLGLYSVWVGKSKDEVNPLD EKIVAKSQELPITNVKQTNGHDVSGAPTNGVVTST"	Db	7083	GTTTTGCTCGGTGTGATCAATGCTGGAGATAATATGTTACTCTGTGGACTTTTGATC	7142
gene	/gene="T12C22.8"	Qy	370	CTTCTCTCTCAACTTCTTCTTGTGATCTCTGCGTCGCAATTGGCTTTTAACGCGCTTC	429
CDS	/note="Contains similarity to an unknown protein F6N23.29 gi 6382042 from Arabidopsis thaliana BAC F6N23 gb AF058919. ESTs gb A1993470, gb AA394704 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAF78264.1" /db_xref="GI:8655991"	Db	7143	CTCTCTGCATCGAGTATTTCGCTCAATTTGCGCTACTCAGTTAGCTTTCAACGGGGTCTC	7202
	/translation="MNKKLLNPLEDPPTPASSDVEDBEISGEDEKHEHISNSSSEEE NELKDLSTQTLNSPTAPTLDSGETNSDSDKPIVLTQSKKEGTDSSGTPKASEET SSDKIKRAKVSQGDNDKQFOSLWKEDEHLSLQCMIDFRAETGTSAHDMWNGFPFIVA KRYTISFDVSKITQFGDKIRGLKKVFKVGRKKGLDLDHDKCLGLAKISLWGLDKEVV LGSSETSNWEKSFYRVVARLVGDVDECIIVKWKSVSKETKKRIEKKMKVNAKELE LLSQKIDLVKEVASVIAETI"	Qy	430	TCCTTACTTCTAAACTCACAAAAAATCACACCACTTATATCTCAATTCACTTGTCTCTTA	489
	/gene="T12C22.9"	Db	7203	TCCTTATTTCAATGCTCAGAAGTTCACCTGCTTGTGATTCCTCACTCCGTGTCTCTTG	7262
	/note="Contains similarity to an unknown protein F6N23.29 gi 6382042 from Arabidopsis thaliana BAC F6N23 gb AF058919. ESTs gb A1993470, gb AA394704 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAF78265.1" /db_xref="GI:8655992"	Qy	490	ACCATAFTCTTACACTTCTTGTATTCCAAATGAACCAAGATCTCCCTCTTCTACTTCA	549
	/translation="MAISPLLWTLIIIFSILSSSSSEEDPTTRFQQYLRFNTHAP NPNTATPISFLNOAQSIGITRTKTFIEISGKPIILLTWLGSNPNLPSILFNSHLDSVP AESBKWTYPFSAHKTIDGHIYARGAODDKICIVQYLESIRNLKSRGFSPLRTIHISY VPEEIGGFDGMWKAFAASPEKDLNLFAMDEGOANPDGDFRVFADRVPMWPHVFIKAE GIPGHGAKLYDNSAMENLMKSVELISRFRESQDFVYKAGRAAYSEVTSNVPYVLYKAGT	Db	7263	TCGTTCTCCGCTGCTTTGTAGTCTCTCAATGATGATGCGGATCTCTCTTCTGGT-----	7316
gene	/gene="T12C22.8"	Qy	550	AAGTCGCGACCAAGTCCAAAGTATGATTTGGATACATCTGCGCGGTGCGGTAGCTCAGCT	609
CDS	/note="Contains similarity to an unknown protein F6N23.29 gi 6382042 from Arabidopsis thaliana BAC F6N23 gb AF058919. ESTs gb A1993470, gb AA394704 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAF78266.1" /db_xref="GI:8655991"	Db	7317	-----GTCTCCAGGTCCTAAGTATATTTGTTGGGTTTGTGTACACTTGTGCGGTCTGCT	7370
	/translation="MNKKLLNPLEDPPTPASSDVEDBEISGEDEKHEHISNSSSEEE NELKDLSTQTLNSPTAPTLDSGETNSDSDKPIVLTQSKKEGTDSSGTPKASEET SSDKIKRAKVSQGDNDKQFOSLWKEDEHLSLQCMIDFRAETGTSAHDMWNGFPFIVA KRYTISFDVSKITQFGDKIRGLKKVFKVGRKKGLDLDHDKCLGLAKISLWGLDKEVV LGSSETSNWEKSFYRVVARLVGDVDECIIVKWKSVSKETKKRIEKKMKVNAKELE LLSQKIDLVKEVASVIAETI"	Qy	610	GTTTATTTCTCTGGTCTTCTTTAAACAGATTACGCGTTTCAAAAAGATTCTAAAGAAATAC	669
	/gene="T12C22.9"	Db	7371	CTCTATTTCTCTGTTGCTATCTTTATGTCAGTTCTGTTTCGAGAGATTTCTGAAGAGGAG	7430
	/note="Contains similarity to an unknown protein F6N23.29 gi 6382042 from Arabidopsis thaliana BAC F6N23 gb AF058919. ESTs gb A1993470, gb AA394704 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAF78267.1" /db_xref="GI:8655992"	Qy	670	ACATTCAAGGCTATTTTATGACATGGCCACATATCCGCTCTATGTTAGTACTTGTGTAGTT	729
	/translation="MAISPLLWTLIIIFSILSSSSSEEDPTTRFQQYLRFNTHAP NPNTATPISFLNOAQSIGITRTKTFIEISGKPIILLTWLGSNPNLPSILFNSHLDSVP AESBKWTYPFSAHKTIDGHIYARGAODDKICIVQYLESIRNLKSRGFSPLRTIHISY VPEEIGGFDGMWKAFAASPEKDLNLFAMDEGOANPDGDFRVFADRVPMWPHVFIKAE GIPGHGAKLYDNSAMENLMKSVELISRFRESQDFVYKAGRAAYSEVTSNVPYVLYKAGT	Db	7431	ACGTTTCTCTGGTCTTCTTGAATGCAATCTACACTTCTTTTAGTGGCGACTTGTGTTTCG	7490
gene	/gene="T12C22.8"	Qy	730	GTGTTAGGACTTTTGGAAAGTGGTGGTGGAAAAGCTGAGTACAGAAATGGAGAGTTT	789
CDS	/note="Contains similarity to an unknown protein F6N23.29 gi 6382042 from Arabidopsis thaliana BAC F6N23 gb AF058919. ESTs gb A1993470, gb AA394704 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAF78268.1" /db_xref="GI:8655991"	Db	7491	GTTATAGGGCTTTTGTCTAGCGGGGAATGGAGAACGCTGCATGGGGAATGGAAAGGTTAT	7550
	/translation="MNKKLLNPLEDPPTPASSDVEDBEISGEDEKHEHISNSSSEEE NELKDLSTQTLNSPTAPTLDSGETNSDSDKPIVLTQSKKEGTDSSGTPKASEET SSDKIKRAKVSQGDNDKQFOSLWKEDEHLSLQCMIDFRAETGTSAHDMWNGFPFIVA KRYTISFDVSKITQFGDKIRGLKKVFKVGRKKGLDLDHDKCLGLAKISLWGLDKEVV LGSSETSNWEKSFYRVVARLVGDVDECIIVKWKSVSKETKKRIEKKMKVNAKELE LLSQKIDLVKEVASVIAETI"	Qy	790	CAACTAGGGAAAAGCTCATACATTTTGTATAAAACATCGGTTTCAACGATATCATGGCAAGCT	849
	/gene="T12C22.9"	Db	7551	CATAAAGGGCAAGCTCTTATGACTGACCTTGGTCTGGACAGCAGTTACTTGGCAAGTG	7610
gene	/gene="T12C22.8"	Qy	850	TGTTTGAATGGAAAGTGTGTTTGAATPATCGAAGTTTTCATCGCTTTTTTCCAAATGTGATA	909
CDS	/note="Contains similarity to an unknown protein F6N23.29 gi 6382042 from Arabidopsis thaliana BAC F6N23 gb AF058919. ESTs gb A1993470, gb AA394704 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAF78269.1" /db_xref="GI:8655992"	Db	7611	TGTTCTGTTGGAGTCGTGGGTTTGTATATTTCTGTTGACGTCGCTCTTCTCAAAAGCTCAT	7670
	/translation="MAISPLLWTLIIIFSILSSSSSEEDPTTRFQQYLRFNTHAP NPNTATPISFLNOAQSIGITRTKTFIEISGKPIILLTWLGSNPNLPSILFNSHLDSVP AESBKWTYPFSAHKTIDGHIYARGAODDKICIVQYLESIRNLKSRGFSPLRTIHISY VPEEIGGFDGMWKAFAASPEKDLNLFAMDEGOANPDGDFRVFADRVPMWPHVFIKAE GIPGHGAKLYDNSAMENLMKSVELISRFRESQDFVYKAGRAAYSEVTSNVPYVLYKAGT	Qy	910	AGCACTCTTTTGTATACAGTTTGCCTGTTCTTTCGCTGTTCTTCTTCTCCGTTGATGAGATG	969
	/gene="T12C22.9"	Db	7671	AGTACGCTCTCTCTAGCTGTGACTCCAATGCTGCTTGGTGTGTTTCCGTTGATAAAATG	7730
	/note="Contains similarity to an unknown protein F6N23.29 gi 6382042 from Arabidopsis thaliana BAC F6N23 gb AF058919. ESTs gb A1993470, gb AA394704 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAF78270.1" /db_xref="GI:8655991"	Qy	970	AGTGAATCAAGTTGGTTCGAATGTTTTTGGCCATCTGGGGATTTGTTTCTTATGGTTAT	1029
	/translation="MNKKLLNPLEDPPTPASSDVEDBEISGEDEKHEHISNSSSEEE NELKDLSTQTLNSPTAPTLDSGETNSDSDKPIVLTQSKKEGTDSSGTPKASEET SSDKIKRAKVSQGDNDKQFOSLWKEDEHLSLQCMIDFRAETGTSAHDMWNGFPFIVA KRYTISFDVSKITQFGDKIRGLKKVFKVGRKKGLDLDHDKCLGLAKISLWGLDKEVV LGSSETSNWEKSFYRVVARLVGDVDECIIVKWKSVSKETKKRIEKKMKVNAKELE LLSQKIDLVKEVASVIAETI"	Db	7731	AGTGGTGTAAAGATTATGGCAATGCTGATCGCTATTTGGGGTTTCGCTTCTTATGTTTAC	7790
gene	/gene="T12C22.8"	Qy	1030	CAGCATTTATGTCAATGATAGAAAGCCAGAGAACCAAGAGCTTCTCTCA	1079
CDS	/note="Contains similarity to an unknown protein F6N23.29 gi 6382042 from Arabidopsis thaliana BAC F6N23 gb AF058919. ESTs gb A1993470, gb AA394704 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAF78271.1" /db_xref="GI:8655992"	Db	7791	CAGAATCATATTGATGATCTTGAAGTAAAGCAAGCAGCAGCAACAAGCTCA	7840

Search completed: November 1, 2004, 15:44:33
Job time : 5081.95 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:48:03 ; Search time 96.7741 Seconds
(without alignments)
7704.726 Million cell updates/sec

Title: US-09-913-767-2

Perfect score: 1049

Sequence: 1 aagatgaagatgaagacagt.....actgattatgtacttaaaa 1049

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PTUS COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73.6	7.0	7218	1	US-08-232-463-14
2	36.2	3.5	705	4	US-09-270-767-5061
3	36.2	3.5	705	4	US-09-270-767-20343
4	36.2	3.5	1827	4	US-09-270-767-1308
5	36.2	3.5	1827	4	US-09-270-767-16590
6	36	3.4	3969	4	US-09-248-796A-5440
7	35	3.3	282	4	US-09-313-294A-5650
8	35	3.3	7120	4	US-09-380-773-2
9	33.4	3.2	399	4	US-09-621-976-8976
10	33.4	3.2	3348	4	US-09-312-762A-2
11	33.4	3.2	14707	4	US-09-312-762A-3
12	33.4	3.2	23439	4	US-08-956-171E-38
13	33.4	3.2	23439	4	US-08-781-986A-38
14	33.2	3.2	289	3	US-09-007-005-17
15	33.2	3.2	289	3	US-09-244-796-17
16	33	3.1	5227	4	US-09-919-172-79
17	33	3.1	5228	4	US-09-919-039-216
18	33	3.1	15144	3	US-08-458-434A-6
19	32.6	3.1	3190	3	US-08-986-768-4
20	32	3.1	285	2	US-08-630-822A-85
21	32	3.1	285	3	US-09-005-069-85
22	32	3.1	285	3	US-09-171-156A-34
23	32	3.1	285	4	US-09-004-730A-34
24	32	3.1	285	4	US-08-981-799A-34
25	31.8	3.0	358	4	US-09-513-999C-24878
26	31.8	3.0	685	3	US-09-183-266A-16
27	31.8	3.0	996	4	US-09-543-681A-3267

C 28	31.8	3.0	1654	4	US-09-634-238-16	Sequence 16, Appl
C 29	31.8	3.0	4177	3	US-09-023-082A-23	Sequence 23, Appl
C 30	31.8	3.0	4177	4	US-09-248-998-23	Sequence 23, Appl
C 31	31.8	3.0	4177	4	US-09-610-651-23	Sequence 23, Appl
C 32	31.6	3.0	240	3	US-08-559-397A-10	Sequence 10, Appl
C 33	31.6	3.0	276	4	US-09-248-796A-1906	Sequence 13019, A
C 34	31.6	3.0	381	4	US-09-248-796A-1906	Sequence 1906, Ap
C 35	31.6	3.0	1261	3	US-08-961-083-25	Sequence 25, Appl
C 36	31.6	3.0	1261	4	US-09-536-784-25	Sequence 25, Appl
C 37	31.6	3.0	1329	4	US-09-583-110-519	Sequence 519, App
C 38	31.6	3.0	3096	4	US-09-614-221A-452	Sequence 452, App
C 39	31.6	3.0	26385	4	US-08-961-527-3	Sequence 3, Appl1
C 40	31.4	3.0	1101	4	US-09-543-681A-3334	Sequence 3334, Ap
C 41	31.4	3.0	2243	1	US-07-995-657-1	Sequence 1, Appl1
C 42	31.4	3.0	2243	1	US-08-474-587-1	Sequence 1, Appl1
C 43	31.4	3.0	2871	4	US-09-489-847-111	Sequence 111, App
C 44	31.4	3.0	5798	4	US-09-377-285B-33	Sequence 33, Appl
C 45	31.2	3.0	711	4	US-09-107-532A-609	Sequence 609, App

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match

7.0%; Score 73.6; DB 1; Length 7218;

Best Local Similarity 2.7%; Pred.No. 9.6e-14;	
Matches 10; Conservative 234; Mismatches 128; Indels 0; Gaps 0;	
Qy	100 GGTGGCAAAAGGATCTGGTTCCAAAGCTTCCTTCAAACCGTGTGGTTGCCACTCATTTTC 159
Db	1048 GGTGAGGAGCTTGCATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1107
Qy	160 TTCCTCTTCTTATCTTCTCCGCCGTCGTGCTGCCCTGAAGAACCAAGAACGACT 215
Db	1108 YY 1167
Qy	220 CCATTTTTCCTCATGAACCTCCTCTCTTATCCCGCTATCGTGTGTGGTTGCTGCTG 279
Db	1168 YY 1227
Qy	280 GGATTTGACAATTAACCTACTCTTACGGGTAGCTTATATCCCTGTTTCTACTGCGTCT 339
Db	1228 YY 1287
Qy	340 TTGATCATCTCCGGGCAATTAGGCTTCACCTGCTCTCTTGGCATTTTATGGTGAAGCAA 399
Db	1288 YY 1347
Qy	400 AAGTTCACACTTTCACATATAAACGCTATCGTTTGTCTCACTGGTGGCGGTAGTCTTT 459
Db	1348 YY 1407
Qy	460 GCCCTTAACCTCT 471
Db	1408 YYYYYYYYYYYY 1419

```

RESULT 2
US-09-270-767-5061/c
; Sequence 5061, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5061
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-5061

```

	Query Match	3.5%;	Score 36.2;	DB 4;	Length 705;
	Best Local Similarity	46.1%;	Pred. No. 0.099;		
	Matches 160;	Conservative 0;	Mismatches 183;	Indels 4;	Gaps 1;
Qy	119	TTCCAAGCTTCCTTTCAAAACCGTTGGTGTGTCACACATATTTCTCCCTCTTCTCTATCTT	178		
Db	593	TTCTTTTCTTCCCTCCCTCTCTTTTTCCTCTCTTTTCTTTTTTCCCTCTCTCTTTT	534		
Qy	179	TCCTCCGGCTGCTGCTTGCCTTGAAGAACAAAGACGACTCCATTTTTCTCATGAAC	238		
Db	533	TCTCTCAATCTCTTTTCTTTTTTCCCTCTCTCTTTTTTCTCTCCCTTTTTTTCTTTTTT	474		
Qy	239	CTCCTCTCTTTATCCCGCTCATCGTTGTGTGTTTTGCTCGTGGATTTGACAAT---	294	TAC	
Db	473	CTCTCTCTTTTTTCTCTCTCTCTTTTCTTTTTTCCCTCTCTCTTTTCTCTCTCTCT	414		
Qy	295	CTCTACTCTTACGGGTAGCTTATACCTCTGTTTCTACTCGCTCTTTGATCATCTCCGG	354		
Db	413	TTTTCTTCGTTCCTCTCTCTTTTTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTTTT	354		
Qy	355	CAATTAGGCTTCACTGCTCTCTTTTGCAATTTTTTATGGTGAAGAAAAAGTTCACACCTTTC	414		
Db	353	TCTCTGCTCTCTCTTTTCTTTTTTCTCTCTCTTTTTTCTCTCTCTCTCTCTCTCTCTCT	294		

```

Qy      415  ACTATAAACGCTATCGTTTTGCTCACTGGTGGTCCGTAGTCCTTCC 461
      |         |         |         |         |         |         |
Db      293  CTCCTCCTTTCTTTTTTCCTTCTCTCTCTCTTCTTCTCTCTCCCTTC 247

RESULT 3
US-09-270-767-20343/c
; Sequence 20343, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila m
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20343
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-20343

```

[illegible]

```

RESULT 4
US-09-270-767-1308/c
; Sequence 1308, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila m
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1308
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1308

Query Match      3.5%; Score 36.2; DB 4; Length 1827
Best Local Similarity 49.0%; Pred. No. 0.19;

```

Matches	124;	Conservative	0;	Mismatches	128;	Indels	1;	Gaps	1;
Qy	126	CTTCCTTCAACCGTGGTGTGTCACATCTTCTCCCTCTCTCTCTTATCTTTCTCTCCG	185						
Db	1151	CTTTCTCTCTCCCTTTCGTTTTCTCTCTCTTTCTCTCTCTCTCTCTCTCTCTCTCTTC	1092						
Qy	186	CCGTCGTGCTGGCCCTGAAGAAACGACATCCATTTTCTCATGAACCTCTCTCT	245						
Db	1091	TCCTCTTTTTCCTATTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1032						
Qy	246	CTTTATCGCGCCTATCGTTGTGTGTGCTCGTGGGATTTGACAATTACCTCTACTCTTA	305						
Db	1031	TTTTTTCTTTCCTTT	972						
Qy	306	CGGGTTAGCTTATATCCCTGTTTCTACTCGGCTCTTTGATCATCTCCGGGAATTAGGCTT	365						
Db	971	CTT-TCCCTCTCTCTTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	913						
Qy	366	CACATGCTCTCTTT	378						
Db	912	CTCTCTCTCTCTT	900						

RESULT 5

```

US-09-270-767-16590/C
; Sequence 16590, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16590
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-16590

```

	Query Match	3.5%; Score 36.2; DB 4; Length 1827;
	Best Local Similarity	49.0%; Pred. No. 0.19;
	Matches 124; Conservative 0; Mismatches 128; Indels 1; Gaps 1;	
Qy	126 CTTTCCTCAACCGTGGTTGTCACATCTTTCTCCCTCCTCTCTATCTATTCTCCTCCG	185
Dd		
	1151 CTTTCCCTCTCCCTTGCTTTTCCCTCTCTCTTTCTCTCTCTCTCTCTCTCTCTCT	1092
Qy	186 CGGTGCGTGGTGCCTCAAGAACGAAGAACGACTCCAATTTTTCTCTCATGAACCTCCTCT	245
Dd		
	1091 TCCTCTTTTTCCTAATGATGTCCTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1032
Qy	246 CTTTATCGCGCATCGTTGGTTTGGCTGCTGGGATTTGACAATTAACCTACTACTCTTA	305
Dd		
	1031 TTTTTCCTTTCCTCTTCCCTTCTCTTTCCCTCTCTCTTTTCTTTCCCTCTCTCTCTTT	972
Qy	306 CGGGTTAGCTATAATCCCTGTTTCTACTCGGTCTTTGATCATCTCCGCAATTAGGCTT	365
Dd		
	971 CTT-TCCTCTCTCTTTTCTTTGCCCTCTCTCTTTTCTTTTCTTCCCTCTCTTTTCTTTC	913
Qy	366 CACTGCTCTCTTT	378
Dd		
	912 CTCCTCTCTCTTT	900

RESULT 6

US-09-248-796A-5440/c
; Sequence 5440, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

```

? TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
?
? FILE REFERENCE: 107196.132
?
? CURRENT APPLICATION NUMBER: US/09/248,796A
?
? CURRENT FILING DATE: 1999-02-12
?
? PRIOR APPLICATION NUMBER: US 60/074,725
?
? PRIOR FILING DATE: 1998-02-13
?
? PRIOR APPLICATION NUMBER: US 60/096,409
?
? PRIOR FILING DATE: 1998-08-13
?
? NUMBER OF SEQ ID NOS: 28208
?
? SEQ ID NO 5440
?
? LENGTH: 3969
?
? TYPE: DNA
?
? ORGANISM: Candida albicans
?
? US-09-248-796A-5440

```

[illegible]

RESULT 7

```

US-09-313-294A-5650
; Sequence 5650, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Ialugudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5650
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350473H1
; NAME/KEY: unsure
; LOCATION: 2, 16, 149, 151, 241, 250, 274, 276
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-5650

```

	Query Match	3.3%	Score 35;	DB 4;	Length 282;
	Best Local Similarity	65.8%;	Pred. No. 0.13;		
	Matches	50;	Conservative	0;	Mismatches 26; Indels 0; Gaps 0;
Qy	812	TCTTCTGTGCATCGTCTCTGGTCTCTGGAATATATGTCAGTGTCTGCTTCCGGTCA	CGG	871	
Db	13	TCCTNCTCTCCACCGATCTGTCTGGCGAACCATGSCCAGCCGGTGTGCTCTCTCATGC		72	

```
Qy 872 TGATCTTGGCGGTCAAT 887
Db 73 TACTGCTGGCGGTCTCT 88

RESULT 8
US-09-380-773-2
; Sequence 2, Application US/09380773
; Patent No. 6759219
; GENERAL INFORMATION:
; APPLICANT: Hein, Silke
; APPLICANT: Soehling, Brigitte
; APPLICANT: Gottschalk, Gerhard
; APPLICANT: Steinbuechel, Alexander
; TITLE OF INVENTION: Methods for the Biosynthesis of Polyesters
; FILE REFERENCE: MORT136--- 118899.0136.NPUS00
; CURRENT APPLICATION NUMBER: US/09/380,773
; CURRENT FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: PCI/US97/03994
; PRIOR FILING DATE: 1997-03-03
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 7120
; TYPE: DNA
; ORGANISM: Clostridium kluyveri
US-09-380-773-2

Query Match 3.3%; Score 35; DB 4; Length 7120;
Best Local Similarity 53.2%; Pred. No. 1.2;
Matches 74; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 314 CTTATATCCCTGTTTCTACTGCGTCTTTGATCATCTCCGCGCAATTAGGCTTCACTGCTC 373
Db 1797 CTTATACTAAGCACTACTCGGGGTTATTCTGTACTAATGACGATATTACCATAAC 1856

Qy 374 TCTTTGCAATTTTATGTTGTAAGCAAGTTCACACCTTTCACATATAACGCTATCGTTT 433
Db 1857 CTTTTCGATATTTTCATATATAAAGAAAGATAAAGGATTAACCATAGTTTCTATTATTG 1916

Qy 434 TGCTCACTGGTGGTCCCT 452
Db 1917 TTTCAATGATGGTGATG 1935

RESULT 9
US-09-621-976-8976/c
; Sequence 8976, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8976
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match 3.2%; Score 33.4; DB 4; Length 399;
Best Local Similarity 13.3%; Pred. No. 0.6;
Matches 27; Conservative 102; Mismatches 73; Indels 1; Gaps 1;

Qy 699 GGTATACAGCAGAGAAGAGATTTTAAAGCTTGGAGAGTCTTTGTACTATGTGGTATG 758
Db 260 GSYRMRAGYRSRWSYTSAMWRKKMTCKGRSSWSGRSTGYAYMYKKSWCTSRKMY 201

Qy 759 TGTGTTACGGCCATAATCTGGCAAGCATTTTTTGTGGAGCTATTGCGGTGATCTTCTG 818
Db 200 YKRRKKRRKCTSTRTCTYRGSTYKCAAYYTKGRKKTWTYYYYSYMSMKKTWRMK 141

Qy 819 TGCATCGTCTCTGCTCTCGAATATGATGTCAGTGTCTGCTCCGGTGACGGTGATCTT 878
Db 140 TAYWTGRKWT-RTKWTCTMCWCTTYWAGTMYRYRYRYAKRAKWSKRCWTSTT 82

Qy 879 GGCGGTCAATTTGCTCCAGGAGA 901
Db 81 CYCMKYWAKKWSYVWSMSMMKW 59

RESULT 10
US-09-312-762A-2
; Sequence 2, Application US/09312762A
; Patent No. 6552177
; GENERAL INFORMATION:
; APPLICANT: MIA HOROWITZ ET AL.
; TITLE OF INVENTION: BH DOMAIN CONTAINING GENES AND PROTEINS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
; STREET: 2001 Jefferson Davis Highway, Suite 207
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead, Slimnote-890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; OPERATING SYSTEM: Windows version 3.11
; SOFTWARE: Word for Windows version 2.0 converted to
; SOFTWARE: an ASCII file
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/312,762A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/026,898
; FILING DATE: 20 FEB 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedmam, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 916/10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-5625553
; TELEFAX: 972-3-5625554
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3348
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-312-762A-2

Query Match 3.2%; Score 33.4; DB 4; Length 3348;
Best Local Similarity 49.7%; Pred. No. 2.6;
Matches 85; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy 832 GTCTCTGGAATATGTCAGTGTCTGCTTCCGGTGACGGTGATCTTTGGCGCGTATTCG 891
Db 1226 GTGATGATATGCTGGCCACCATATAGCTCGGTGATGTTGGTGGTGGAGGAG 1285

Qy 892 TTCAGAGAGATTTTCAGCGGGGAAGGTGTCGCTTTGGCTCTCTCCCTCTGGGGATCA 951
Db 1286 TCCTGTATGCCCTCACAGGCTGTGAAGGGTGGTGTCTTTTATGGCACCATGAATGGGCC 1345

Qy 952 GTCTCTTATTTCTATGACAGGTTAAATCCGAGGAGAAGACTAAGGCTCAG 1002
```

Db 1346 TTTGGGCGATGGCTACCGCGAGGGGGCTGGCGAGGGCAATTGATGATGTTGAG 1396

RESULT 11

US-09-312-762A-3

; Sequence 3, Application US/09312762A

; Patent No. 6552177

; GENERAL INFORMATION:

; APPLICANT: MIA HOROWITZ ET AL.

; TITLE OF INVENTION: EH DOMAIN CONTAINING GENES AND PROTEINS

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

; STREET: 2001 Jefferson Davis Highway, Suite 207

; CITY: Arlington

; STATE: Virginia

; COUNTRY: United States of America

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

; COMPUTER: Twinhead* Slimnote-890TX

; OPERATING SYSTEM: MS DOS version 6.2,

; OPERATING SYSTEM: Windows version 3.11

; SOFTWARE: Word for Windows version 2.0 converted to

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09312,762A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/026,898

; FILING DATE: 20 FEB 1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Friedmam, Mark M.

; REGISTRATION NUMBER: 33,883

; REFERENCE/DOCKET NUMBER: 916/10

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 972-3-5625553

; TELEFAX: 972-3-5625554

; TELEX:

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 14707

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-09-312-762A-3

Query Match

Best Local Similarity 3.2%; Score 33.4; DB 4; Length 14707;

Matches 85; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 832 GTCTCTGGAAATATGGTCAGTCTCTGCTTCGGGTGACGGTGATCTTGGCGGTCAATTGC 891

Db 8947 GTGGATGATATGTCGCCCAACGATATACTCGGCTGATGCTGTCGCCCGCAGGAGG 9006

QY 892 TTCCAGGAGAAAGTTTCAGCGGGGAAAGGTGTCGCTTTCGGCTCTCTCCCTCTGGGATCA 951

Db 9007 TCCTGTATGCTCCACAGCTGTGAAGGTGTGCTTTTGATGGCACCAGTAATGGGCC 9066

QY 952 GTCTCTATTTCATGGACAGGTTAAATCCGAGGAGAAAGCTAAGGCTCAG 1002

Db 9067 TTTGGCATGGCTACGGCGAGGGGCTGGCGAGGGCAATTGATGATGTTGAG 9117

RESULT 12

US-08-956-171E-38

; Sequence 38, Application US/0895617E

; Patent No. 6591114

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; Gil H. Choi

; Patrick S. Dillon

; Craig A. Rosen

; Steven C. Barash

; Michael R. Fannou

; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 5256

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/956,171E

; FILING DATE: 20-Oct-1997

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/009,861

; FILING DATE: January 5, 1996

; APPLICATION NUMBER: 08/781,986

; FILING DATE: January 3, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Mark J. Hyman

; REGISTRATION NUMBER: 46,789

; REFERENCE/DOCKET NUMBER: PB248P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (240) 314-1224

; TELEFAX: (301) 309-8439

; INFORMATION FOR SEQ ID NO: 38:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 23439 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 38:

US-08-956-171E-38

Query Match 3.2%; Score 33.4; DB 4; Length 23439;

Best Local Similarity 60.4%; Pred. No. 9.6;

Matches 55; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 688 GCGGATTTCAAGGTGATAGCAGGAGACCAAGAGATTTTAAGCTTGGAGAGCTCTTTGTAC 747

Db 13437 GGTGATTATATGTTAAGAGAGCAAGAGAAAGATAGTATTAGAAAGTATTCAATA 13496

QY 748 TATGTGGTGATTGTGTTCCACGCCATAATCT 778

Db 13497 GCGTGGTGTCAGTGTAGCGGCTACAATGT 13527

RESULT 13

US-08-781-986A-38

; Sequence 38, Application US/08781986A

; Patent No. 6737248

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 5255

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

[illegible]

Search completed: November 1, 2004, 22:40:40
Job time : 104.774 secs

THIS PAGE BLANK (USPTO)

1	337.6	32.2	1233	16	US-10-424-599-30274	Sequence 30274, A
2	316.2	30.1	1146	16	US-10-425-114-10516	Sequence 10516, A
3	316	30.1	994	16	US-10-424-599-73501	Sequence 73501, A
4	231.4	22.1	1173	17	US-10-437-963-57539	Sequence 57539, A
5	167.6	16.0	3387	9	US-09-938-842A-2212	Sequence 2212, Ap
6	167.6	16.0	3387	11	US-09-938-842A-2212	Sequence 2212, Ap
7	165.2	15.7	2214	16	US-10-424-599-102519	Sequence 102519, A
8	151	14.4	1513	17	US-10-437-963-82061	Sequence 82061, A
9	144.4	13.8	1296	16	US-10-425-114-13687	Sequence 13687, A
10	142	13.5	1177	17	US-10-437-963-91425	Sequence 91425, A
11	133	12.7	83698	17	US-10-416-898-9	Sequence 9, Appli
12	130.4	12.4	789	17	US-10-437-963-72452	Sequence 72452, A

13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

Qy	200	TTGAAGAACAAAGACGACTCAATTTTTCCTCATGAAACCTCCTCTCTTTATTCGCCGCTA	259
Db	122	ATCGCT-----CTTCCGCCCTAAATCTCTATCAAGCCCCCTCTCTTCTCGCTCGCTCGG	172
Qy	260	TCGTTGTGTGGTTGCTCGTGGGATTTGACAAATTACCTCTACTCTTCACTTACGGGTAGCTTATA	319
Db	173	CCCTCATCGGACTCTCTACCGGCTCGACGACTACCTCTACGCTGCGGCGTGGCTCGCC	232
Qy	320	TCCCTGTTTCTACTGCGCTTTTGATCATCTCCGCGCAATTAGGCTTCACTGCTCTCTTTG	379
Db	233	TTCCGGTCTCCACTTCTCTCTAATCCAAGCCTCCACCTCGCTTCAACCGCGCTTCG	292
Qy	380	CAITTTTATTCGTGAAGCAAAAGTTACACCTTTACACTATATAACGTAATCGTTTTCGTCA	439
Db	293	CTTTCTCTCTCGTCCGCACAGGTTACGCGCTACTCCGTCAACTCCGTGCTGCTTCTCA	352
Qy	440	CTGCTGTGGCCGTAGTCTCTGCCCTTAACTCTGATAGTGAACAAGCTTCAAAACGAGACAC	499
Db	353	CCGTGCGGCTGTGGTTCTGGCTCTGGCTTCAGCGGGACCGCCCGCGGTGAGTCTGA	412
Qy	500	ACAAGGAATATGTTGTTGGGTTCAATCATGACTCTTTGGTGCAGCTCTTCTCTATGGTTTA	559
Db	413	GTGCCAGTACGTGATGGTTTGTATGATACTTTGGCTGCGCGCTGTATGGTTTCG	472
Qy	560	TATTTGCCACTTGTGAGGTTCTTACAAGAAATCTGGTCAGGAATCAAGTATACGCTCG	619
Db	473	TTTTGGCGGTGATGGAGTTGGTGACAAAAAGACAGCAGCGTATCACGTACTCTCTGG	532
Qy	620	CGCTCGAGTTCCAGATGGCTTATATGCTTTGTCTGTCACCTTGTCTGCTCGTGGGGATGC	679
Db	533	TCATGGAGATTCAGCTTGTCTTGTGCTCTTTGTACCTATTCTGACCCGTTGGAATGA	592
Qy	680	TAGCCGCTGGCGATTTCAAGGTGATAGCAGGAGAAGCAAGAGATTTTAAGCTTGGAGAGT	739
Db	593	TAAATCAATATGACTTCAAGGTGATTCGCGAGAAGCAAGAGATTTTAAGCTTGGGAAA	652
Qy	740	CTTTGTACTATGTGGTGAATGTGTACGGCCATAATCTGCGCAAGCATTTTGTGGGAG	799
Db	653	CAAAAGTACTACGTTGTGTGTGGAGTGCAATAATGTGGCAGTTTTTCTTTGGGAG	712
Qy	800	CTATGGGTTGATCTTCTGTGCATCGTCTCTGGTCTCTGGAATTATGGTCAGTGTCTGTC	859
Db	713	CAATAGGGTTATCTTTTGTGCCCTCGTCTTGTGTGTCGGTATTATAATTGCTGCTTTTC	772
Qy	860	TTCCGGTGAACGGTGATCTTGGCCGTCATTTCGTTCCAGAGAGAAGTTTCAGGCGGGAAAG	919
Db	773	TTCAGTGAACGGAAGTTTGGCTGTATTGTATACAAAGAGAGAGCTTTTCATGCAGAGAAG	832
Qy	920	GTGTGCTTTTGGCTCTCTCCCTCTGGGGATCAGTCTCTATTATTCTATGCAGCAGGTTAAAT	979
Db	833	GGGTTGCTTTTGGTGTCTCTCTTTTGGGGGTTTGTCTCTATTCTATGAGAGATAAATAC	892
Qy	980	CCGAGGAGAAGACTAA	995
Db	893	AAGACAGGGAAGAA	908

RESULT 2

US-10-425-114-10516
; Sequence 10516, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei

```

; AFFILIATE: Cals, Longwell
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
;

```

```

? NUMBER OF SEQ ID NOS: 73128
?
? SEQ ID NO 10516
? LENGTH: 1146
? TYPE: DNA
? ORGANISM: Glycine max
?
? FEATURE:
?
? OTHER INFORMATION: Clone ID: 700941791_F11
US-10-425-114-10516

```

Query Match 30.1%; Score 316.2; DB 16; Length 1146;
Best Local Similarity 63.2%; Pred. No. 5.6e-90;
Matches 486; Conservative 0; Mismatches 283; Indels 0;

Qy	227	TCCTCATGAAACCTCCTCTCTTTATCGCCGCTATCGTTGTGGTTTGTCTCGTGGGAATTC	286
Db	61	TCTCTATCAAGCCCCCTCTCTCTCGCTCGCCCTCATCGGACTCTCTCACCGCCCTCG	120
Qy	287	ACAATTACCTCTACTCTTTACGGGTTAGCTTATATCCCTGTTTCTACTCGCTCTTTGATCA	346
Db	121	ACGACTACCTCTACGCCTGCGCGTGGCTCGCCTTCGCTTCGCTCTCACTTTCTCTCTAATCC	180
Qy	347	TCTCCGGCGCAATYAGGCTTCACCTGCTCTCTTTTGCAATTTTTATGGTGAAGCAAAAGTTCA	406
Db	181	AAGCCTCCCACTCGCCTTCACCGCGCTTCCTCGCTCCGCCACAGGTTCA	240
Qy	407	CACCTTCACTATAAAACGCTATCGTTTGTCTACTGTGTGGTGGCGTAGTCTTGTCCTTA	466
Db	241	CGCCCTACTCCGTCAACTCCGTCGTGTCTCTCACCGTCGCCGCTGTGGTCTCTGGCTCTGC	300
Qy	467	ACTCTGATAGTACAAGCTTGCAAACGAGACACACAAGGAATATGTTGTGGTTTCATCA	526
Db	301	GTTCCAGCGGGGACCGGCCCGCGTGAGTCGAGTCGCCAGTACGTGAATGGTTTGTTA	360
Qy	527	TGACTCTTGTGTGACGCTCTTCTCTATGSGTTTATAATTGCCACTTGTTCGAGCTTTCCTTACA	586
Db	361	TGATACTTGGCGCTCGCGCGCTGATGGGTTTCGTTTTCGCGTGTGATGAGTTGGTGTACA	420
Qy	587	AGAAATCTGTCAGCGAATACGATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCT	646
Db	421	AAAAGAGCAGCAGCGATCACGTACTCTCTGGTCATGGAGATTCAAGTGTCTGTGCT	480
Qy	647	TTGCTGCCACTGTGTCTGCTCGCTCGTGGGATGCTAGCCGCTGGCGATTCAAGGTGATAG	706
Db	481	TCTTTGCTACCTTATCTCGACCGTTGGAATGATAATCAATATGACTTCAAGGTGATTC	540
Qy	707	CAGGAGAAGCAAGAGATTTTAAAGCTTGGAGAGTCTTTGTACTATGTGTGATTTGTCTTCA	766
Db	541	CGCGAGAAGCAAGAGATTTTAAAGCTTGGGGAACAAAGTACTACGTTGTGTGGTGTGGA	600
Qy	767	CGGCCATAATCTGCCAAGCATTTTTTGTGGAGCTATTTGGGTTGATCTCTGTGATCGT	826
Db	601	GTGCAATAATGTGCGAGTTTTTCTCTTTGGGAGCAATAGGGGTTATCTTTTGTGCTCGT	660
Qy	827	CTCTGGTCTCTGGAATATGCTCAGTCTCTGTTTCCGCTGACGCTGATCTTTGGCGCTCA	886
Db	661	CTTTTGTGTCGGTATTATAAATGCTGCTTTTCTTCCAGTGACGGAAGTTTGGCTGTTA	720
Qy	887	TTTTGCTTCCAGGAGAAGTTTCAGCGGGGAAAGGTGTGCTTTTGGCTCTCTCCCTCTGGG	946
Db	721	TTGTATACAAAGAGAGCTTTCATCGCAGAAAGGGTTGCTTTGGTGTCTCTCTTTGGG	780
Qy	947	GATCAGTCTCTTATTTCTATGGA CAGGTTTAAATCCGAGGAGAAGACTAA	995
Db	781	GGTTGTGTCTCTATTTCTATCGAGAGATAAAACCAAGA CAGGGAAGAAGAA829	

RESULT 3

US-10-424-599-73501
; Sequence 73501, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 73501
LENGTH: 994
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_37387C.1
US-10-424-599-73501

Query Match 30.1%; Score 316; DB 16; Length 994;
Best Local Similarity 61.6%; Pred. No. 6e-90;
Matches 541; Conservative 0; Mismatches 330; Indels 7; Gaps 2;
QY 10 ATGAAGACAGTTCCTGTAATCATAAACTGTATATCTTGGCCATTGGAACCTGTGGAGC 69
DB 118 ATGAAGCGCTTCTCTCAGCAAACTGTCTTACTCACCATCGGCACTCCGGTGG 177
QY 70 CCTCTAATGATGCTCTTACTTCCAAATGGTGGCGAAGATCTGTGTTCCAACTTC 129
DB 178 CCCCCTCGTATGCTCTTACTTCTCCACGGCGGCCACCGGCTCTGGCTCTCCAGCTTC 237
QY 130 CTTCAACCGTGGTGTCCATCATTTTCTCCCTCTTCTTATCTTCTTCTCCCGCT 189
DB 238 CTCGAACCGTGGCTTCCCTCTCATGCTCTCCCTCCCTCGGCTCTCTTACCCTCGCGCA 297
QY 190 CGTCG-----TTGCTTGAAGAACAAAGAACGACTCCATTTTCTCTCATGAACTCTCT 243
DB 298 CGTCGACCGCTCCGCGCCGCGAACCGCAACCAATTAATCTCAATGAAGCTCTCT 357
QY 244 CTCCTTATCGCGCTATCGTTGTTGTTGCTCGTGGGATTTGACAAATPACTCTTACTCT 303
DB 358 CTCCTCGCGCTCCCACTTCTATCGGAATCTCACGGGCTCGAGACTTACCTCTACGCC 417
QY 304 TAGCGGTAGCTTATATCCCTGTTTACTGCGTCTTGTGATCATCTCCGCGCAATTAGGC 363
DB 418 TAGCGGTGGCAGGCTTCCGCTCTCACTTCGCGCTCATCATCGCAACGCAACTCGGC 477
QY 364 TTCACCTGCTCTCTTTCATTTTATAGTGAAGCAAAAGTTCACACCTTTTCACTATAAC 423
DB 478 TTCACGCGTTCCTTCGCTTCCTCTCGTAGCAGAAAGTTACGCGGTACTCCGTAAC 537
QY 424 GCTATCGTTTGTCTCACTGGTGGTGGCGTTCCTTGGCCCTTAACTCTGATAGTACAAAG 483
DB 538 GCGCTGTTTGTCTCACTGTGCGCGCGCGTTTTGGCGCTTTCACACACGCGGAGACCGT 597
QY 484 CTTGCAACGACACACAGAAGAAATATGTTGTTGGTTTCATCATGACTCTTGGTGCAGCT 543
DB 598 CCCCCTGGCGAGTCCGTTAAGGAATATGTTATGGGCTTTGTGATGACAGTATCGCTGCG 657
QY 544 CTTCTCTATGGGTTTATATGCACTTGTGAGCTTTCTTACAGAAATCTGGTCAGCGA 603
DB 658 GCATTGTATGATTCATTTTACCTTGGTGGAGTTGGGTACAAAATATCAACAGCCT 717
QY 604 ATCAGTATACCTCGCGCTCGAGTTCCAGATGGTCTTATGCTTTGCTGCACTTGTGTC 663
DB 718 CTTACTTACTCTCTTGTGATGAGATTCAAGTTCGTTATGTTCTTCTCGGCCACTCTCTT 777
QY 664 TGCCTCGTGGGATGCTAGCCCTGGCGATTTCAAGGTGATAGCAGGAGAACAGAGAT 723
DB 778 TGCCTCTTGGAAATGATCATCAACAAATGACTTTAAGGTGATTCGAGGGAAGCCAAAAA 837
QY 724 TTTAAGCTTGGAGAGTCTTTGTACTATGTTGGTGAATTTGTTTCCAGCGCCATATCTGGCAA 783
DB 838 TTTGAGCAGGAGAGGAAGTTACTATGCTGTTTGGTGGGAGTGCATAATATTATGGCAG 897
QY 784 GCATTTTGTGGGAGCTATTGGGTTGATCTTCTGTGCACTGCTCTGTGCTCTGGAATT 843

DB 898 GCTTTTCTTGGGCGCATTTGGGTTATATTTTGGGCTGGCTTCTGATCT-AGGATTT 956
QY 844 ATGGTCAGTCTCTCTTCCCGTGACGGTGATCTTTGGC 881
DB 957 TTGATTGCGGTGTTGCTATACCCGTAACCGAAGTGTGGC 994
RESULT 4
US-10-437-963-57539
Sequence 57539, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 57539
LENGTH: 1173
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_59340C.1
US-10-437-963-57539

Query Match 22.1%; Score 231.4; DB 17; Length 1173;
Best Local Similarity 54.1%; Pred. No. 8.1e-63;
Matches 540; Conservative 0; Mismatches 426; Indels 33; Gaps 2;
QY 9 GATGAAGACAGTCTTGTAAATCATAACTGTATATCTTGGCCATTGGAACTGTGGAGG 68
DB 102 GTTCAGGAGCCGCTCTCGTCACTTCGTCCTCATGTTGGTGGCTCGGCTGCGG 161
QY 69 CCCTCTAATGATGCTCTCTACTTCCAAAATGGTGGCGAAAGATCTGGTTTCCAAGCTT 128
DB 162 GCGCTCTCTCTCGGCGCTACTTCTCGCGGCGGCAACCGCAAGTGGCTCTCCAGCT 221
QY 129 CTTCAAAACGTTGGTGTCCACTCATTTTCTTCCCTCTCTCTTATCTTCTCCGCG 188
DB 222 TCTCAGACCGCGCTGGCGCTGCTGCTCGCGCGCTCTGCTTCTGCTACTCTCAG 281
QY 189 TCGCTGTTGCTTGAAGAACAGAAA-----CGACTCCATT 224
DB 282 CCGCGCGCGCGAGGTTGAGGACGACGCGCTGGCGCTGGCGCGGCGCACCGCT 341
QY 225 TTTCCTCATGAAACCTCTCTTTATCGCGCTATCGTTGGTTGGTTGCTGCTGGGATT 284
DB 342 GTTCTCATGAGCCACGCTCTCGTGGTGGCTCGCGCTCGTGGGCTCATGACCGGCT 401
QY 285 TGACAAATPACTTACTTCTTACGGGTAGCTTATATCCCTGTCTTCTACTGCGTCTTTGAT 344
DB 402 CGAGACCTCTTACGCTACGCTACGCGCTACCTCCCGGTGTCCACTCTCTCATCT 461
QY 345 CATCTCCGCGCAATTAGGCTTCACTGCTCTTTGTCATTTTATGTTGAAGCAAAAGTT 404
DB 462 CATCTCAGCAGCTGGCTTTCACGGCGGCTTGGCTGCTGCTGCTGCTGCGCGAGCGGT 521
QY 405 CACACTTCTACTATAAACGCTATCTGTTTGTCTCACTGCTGCTGCGGTAGTCTTGGCCT 464
DB 522 CACGGCTTCTGGTGAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 581
QY 465 TAACCTCTGATAGTGAACGCTTGAACGAGACACACAGGAATATGTTGTTGGTTCAT 524

Db 582 GAACGCGCGGGGACCGCGCGGGGTGTGCGCGCGCAGTACTGCGCGGGTTGCG 641
Qy 525 CATGACTTTGGTGGAGCTCTTCTATGCGGTATATATGCACTTGTGAGCTTTCTTA 584
Db 642 CATGACGCTCGCGCGGGCGCTGTATGCGGCTGTGCTCCCGCTCATGAGCTCAGCA 701
Qy 595 CAAGAA-----ATCTGTGTCAGCAATCAGCTATACGCTCGCGCTCGAGTTCCAGAT 635
Db 702 GCGGCACACGCGCGCGCGCGCGCTTACCGTCACTGCTCGTATGAGATGCGAGCT 761
Qy 636 GGTCTTATGCTTTGTGCACTTGTGCTGCTCGTGGGATGCTAGCGCTGCGGATTT 695
Db 762 CGTATCGCTTGTGCGCACCGCTTACGCGCGTGGCATGCTCGTCAACAGGATTT 821
Qy 696 CAAGGTATAGCAGAGAGAAAGATTTTAAAGCTTGGAGAGTCTTTGTACTATGTGT 755
Db 822 CCAGGCAATCCAGAGAGAGCCATGAATTCGGTCTGGGCCCAAGCGGCTACTACCTGCT 881
Qy 756 GATTGTGTTACGGCCATAATCTGGCAGCATTTTGTGGGAGCTATGGGTTGATCTT 815
Db 882 CTGCGCGGATCGCGCGCATGTACAGTCTTCTTCGCGCAGATCGCGCGCATCTT 941
Qy 816 CTGTGATCGTCTCTGGTCTCTGGAATTAATGCTCAGTCTCTGCTTCGCGTGAAGGTGAT 875
Db 942 CTACGGCTCGCGCTGCTCGCGCGTCACTATGACCGTGTCTATCCGCTCACCGAGT 1001
Qy 876 CTGCGCTCATTTGCTTCCAGAGAGATTTTCAAGCGGGGAAAGTGTGCTTTGGCTCT 935
Db 1002 GCTCGCGCTCATTTTCCAGAGCCATTAACGGCACAAAGGCGCTCGGCTCGCTCT 1061
Qy 936 CTCCTCTGCGGATCAGTCTCTTATTTCTATGGACAGT 974
Db 1062 CTCGCTATGGGCTTCGTCTCTACTTCTACGGCGAGGT 1100

RESULT 5

US-09-938-842A-2212
; Sequence 2212, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT FILING DATE: 2001-08-24
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2212
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2212

Query Match 16.0%; Score 167.6; DB 9; Length 3387;
Best Local Similarity 48.6%; Pred. No. 4e-42;
Matches 492; Conservative 0; Mismatches 514; Indels 6; Gaps 1;
Qy 1 AAGTGAAGATCAAGACAGTCTTCTTAATCATATAAATCTATATCTTGGCCATTTGGAAC 60
Db 2353 AAGAACTGTAGAGTGGCTCGTGTCTCCATATACGCAATCTTGTATCTTCTGCA 2412
Qy 61 TGTGAGGCCCTCTAATGATCGTCTCTACTTCCAAAATGTGGCGAAAGGATCTGGTTT 120
Db 2413 CCACCTTGCTACAGTTCTGGTAGACTGTACTATGAATGAGGAAAGGACATATGTG 2472

Qy 121 CCAGCTTCTCTCAACCGTTGGTGTGCACATATTTCTTCCCTCTTCTCTTATCTTTC 180
Db 2473 GTAACACTTCTCAACTCATTTGGCTTCCCTGTGTAATCTCTTCCGCTCTTCTTCTCGA 2532
Qy 181 CTCCGCGCTCGTGTGGCTTGAAGAACAAACAAAGCACTCCATTTTCTCATGAACCT 240
Db 2533 ATCAGGCAACCAATCAACAGATACAAATTTGAGTGTCCCTTCTTCCACCCCTT 2592
Qy 241 CCTCTCTTTATCCGCGCTATCGTGTGTGGTGTCTCGTGGGATTTGA CAATTAACCTCTAC 300
Db 2593 GCATCGGTTTAC-----TTGTGCACTGGAGCTGTAGTCTGTCTATGCTTATTTGTCT 2646
Qy 301 TCTTACGGGTAGCTTATATCCCTGTTTCTACTGCGTCTTTGATCATCTCGGCGCAATTA 360
Db 2647 GCAGTTGGGTGTGTTTACTTTACAGTCTCTACTTTCTCCCTCATCTTGGGCTCACAAGTTG 2706
Qy 361 GGCTTCACTGCTCTCTTTGCAATTTTATGGTGAAGCAAAAGTTTCAACCTTTTCACTATA 420
Db 2707 GCCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTTCACTCTTGTAGTGC 2766
Qy 421 AAGCTATCGTTTGTCTACTGTGTGGTGGTGTCTTGGCTTAACTCTGTAGTAGTGC 480
Db 2767 AATTCTTTGTTTCTCTTACGGTTTCTCTGCGCTCTCTCGTGGTCAACACTGATTCAGAA 2826
Qy 481 AAGCTTCAACAGACACACAAGGATATCTTGTGGGTTTCACTCATGACTCTTGTGTGCA 540
Db 2827 AACACAAATGTATCTAGAGTACAAATATGTGATGGGTTTCACTGTACATTTGGTGTCT 2886
Qy 541 GCTCTTCTATGGGTTTATATTTGCCACTTGTGTGAGCTTTTCTTACAGAAATCTGTGTGAG 600
Db 2887 TCCGCTGGGATTTGGATTTGTACTATCTCTGATACAACTGTCTTTCAGAAAGTTTTCACG 2946
Qy 601 CGAATACGATATACGCTCGCGCTCGAGTTCCAGATGCTTATGCTTTGTGTGCCACTTGT 660
Db 2947 AAGCATACATCTCAGCAGTCTCTGGACTTGGCCAAATACCAAGTCTCTAGTTGCAACTGT 3006
Qy 661 GTCTGCTCGTGGGATGCTAGCGCTGGGATTTCAAGGTGATAGCAGGAGAAAGCAAGA 720
Db 3007 GTGGTACTCATAGACTGTTTGAAGTGGAGAGTGGAGAACTCTGCCAAGTAGATGAGA 3066
Qy 721 GATTTTAAGCTTGGAGAGTCTTTGTACTATGTGTGATTTGTGTTCAGCGCATATCTGG 780
Db 3067 AACTACAACTGGGAAAGTGCATATATCTTGACTTTGGCTCAGCAGCTATTTTCTGG 3126
Qy 781 CAGCAATTTTGTGGGAGCTATGGGTGATCTTCTGTGCATGCTCTGTGGTCTCTGGA 840
Db 3127 CAAGTATACACTGTTGGTGTGTGGGATTAATCTTCGAGTCTTCTTCTGTGTCTCCAAT 3186
Qy 841 ATTATGCTCAGTGTCTGTCTTCCGGTGACGCTGATCTTGGCGCTCATTTGCTTCCAGGAG 900
Db 3187 TCCATAACTGCTGTGGGACTGCTTATAGTTCCAGTTGTAGCAGTAGATTTTCCATGAT 3246
Qy 901 AAGTTTTCAGGGGGAAGGTGTGCTTGGCTCTCTCCCTCTGGGATCAGTCTCTTAT 960
Db 3247 AAGATGATGATCAAAAGATTTTCTCCATCATTTTACGATCTGCGGCTTCTTCTCATTT 3306
Qy 961 TTTATGGAAGGTTAAATCCGAGGAGAAAGTAAAGCTCAGGATACACAA 1012
Db 3307 GTCTATCAGCACTACTCCGACGAAAGAGTTGAAGACTTGGCCAGACAAAAC 3358

RESULT 6

US-09-938-842A-2212
; Sequence 2212, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE

```
; FILE REFERENCE: SRR1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2212
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2212

Query Match      16.0%; Score 167.6; DB 11; Length 3387;
Best Local Similarity 48.6%; Pred. No. 4e-42;
Matches 492; Conservative 0; Mismatches 514; Indels 6; Gaps 1;

QY 1 AAGATGAAGATGAAGACAGTCTTGTAAATCAATAAAGTATATCTTGGCCATTGGAAAC 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2353 AAGAACTGTAAAGGTGGCTCGGTCTCCATATAGCCTATCTTGTCTCTCGCAA 2412
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 TGTGGAGCCCTCTAATGATCGCTCTCTACTTCCAAATGGTGGGAAAGGATCTGGTTT 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2413 CCACCTTGTACAGTCTCTGGGTAGACTGTACTATGAAATGGAGGAAAGCACAATATGT 2472
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 121 CCAAGCTTCTTCAACCCGTGGTGTCCACTCATTTTCTCCCTCTCTCTTATCTTTC 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2473 GTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGATCTCTTCGGCTTCTTTCTCGA 2532
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 181 CTCGGCGCTGCTGTGCTTGAAGAAACAAGAAAGCACTCCATTTTCTCATGAACCT 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2533 ATCAGCAACCCAAATCAACAGATACAAATTCAGTCACTGCCCTTCTTCAACCCCTT 2592
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 241 CCTCTCTTTATCCCGCTATCGTGTGTTGCTCGTGGGATTTGACAAATTAACCTCTAC 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2593 GCATCGGTTTAC-----TTGTGCACTGGACTGTAGTGTCTGCTTATTTGTCT 2646
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 301 TCTTACGGGTAGCTTATATCCCTGTTTCTACTGGCTTTTGATCATCTCCCGCAATTA 360
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2647 GCAGTGGGTGCTTTACTTACCAGTCTCTACTTTCTCCCTCATCTTGGCCCTCACAGTTG 2706
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 361 GGTTCACCTGCTCTCTTTCATTTTATGGTGAAGCAAAAGTTCACACCTTTTCACTATA 420
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2707 GCTTCACCTGCTTTTCTATATTCCTTAATCTGCAAAAGTTCACCTTTTGTATGTC 2766
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 421 AACGCTATCGTTTGTCTCACTGGTGGTGGCGTAGTCTTTCGCCCTTAACTCTGATAGTAC 480
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2767 AATTCTTTTCTCTTACGGTTCCTCTGCCCTCTCGTGGTCAACACTGATTCAGAA 2826
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 481 AAGCTTGAACAGACACACAAAGAAATATGTTGGTTCATCATGACTCTTGGTGA 540
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2827 AACAACAACAAATGTATCTAGAGTACAATATGTGTTGGTTCATCTGTACCAATGGTGTCT 2886
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 541 GCTCTTCTATGGGTTTATATGACCTTGTGCGAGCTTCTTCAAGAAATCTGTCAG 600
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2887 TCCGCTGGGATGGATGGTGTACTATCTCTGATACACTGCTCTTCAGGAAGTTTTCAG 2946
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 601 CGAATCACATATACGCTCGCGCTCGAGTTCCAGATGGTCTTATGCTTTGCTGCCACTTGT 660
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2947 AAGCATACATCTCAGCAGCTCTGGACTTGGCCAAATACCACTCTAGTTGCACTTGT 3006
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 661 GTCTGCTCGTGGGATGTAGCCGTGGCGATTTCAAGGTGATAGCAGGAGAAAGCAAGA 720
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3007 GTGTGACTCATAGGACTGTTTGAAGTGGAGAGTGGAGAACTCTGCAAGTGGATGAGA 3066
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 721 GATTTTACGCTTGGAGAGCTTTTGTATGATGTTGTTGTTTCAACGCCCATATCTGG 780
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3067 AACTACAAACTGGGAAAGTGTATATATCTTTGACTTTGGCTCTCAGCAGCTATTTCTGG 3126
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 781 CAAGCATTTTTTGTGGAGCTATTGGGTTGATCTTCTGTGCACTCGTCTCTGCTCTGGA 840
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 7
US-10-424-599-102519
; Sequence 102519, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 102519
; LENGTH: 2214
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63592C.1
US-10-424-599-102519
```

```
Query Match      15.7%; Score 165.2; DB 16; Length 2214;
Best Local Similarity 51.8%; Pred. No. 1.9e-41;
Matches 374; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

QY 245 TCTTATCGCCGCTATCGTGTGTTTGTCTGCGGATTTGACAAATTAACCTTACTCTT 304
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 623 TCTTTTGTATATATTTTGTCTTTGGAGTCTTAATGCTGTCACAAATATGATGACTCA 682
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 305 ACGGTTAGCTTATATCCCTGTTTCTACTGCGTCTTTGATCATCTCCGCAATTAGGCT 364
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 683 CTGGACTCTTATACCTCTCGGCTTCTACCTATTGCTGATTTGTGTCATCAGATTAGCTT 742
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 365 TCACCTGCTCTTTTGGCAATTTTATGTTGAAGAAAGTTCACACCTTTCATATAAAG 424
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 743 TTAATGCAAGTTTCTCATATTTTATCAATTTCTCAAAAGTTTCACTGCCCTTGAATAAAT 802
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 425 CTATCGTTTTCCTCACTGCTGCGTAGTCTTTCGCCCTTAACTCTGATAGTGAAGC 484
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 803 CTACAGTGGTCTTCTCATTTTATCTGCTGCTGCTCTTCTGTTGTTAAGCAAGACAGATGAAC 862
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 485 TTGCAACGAGACACACAAGGAATATGTTTGGGTTTCATCATGACTCTTGGTGCAGCTC 544
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 863 CATCTGGTTTCTCAAGGAAAGTACATATTGTTTTCCTATGTACCTTGGAGCTTCTG 922
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 545 TTCTTATGGGTTTATATATGCACTTGTGAGCTTTTCTTACAAGAAATCTGGTCAAGCAA 604
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 923 CAGTGTACTCTCTTTTGTCTTCCCTCACTGAGCTGACCTTTTGAGAAGGTTCTGAGAAGG 982
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 605 TCAGGTATACCTCGCGCTCGAGTTCCAGATGGTCTTATGCTTTTGTGCCACTTGTGTCT 664
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 983 AACAATTTCTGTTGTTTGGAAATGCAAAATCTACACATCAATCTGTTGCCCTCTGTTGCTT 1042
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 665 GCCTCTGGGATGCTAGCCGCTGCGGATTTCAAGGTTGATAGCAGGAGAAAGCAAGATTT 724
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 1043 CTGTCATAGGCCTATTTCGAAGTGGGAAATGGCGTACTTTTCATGAGCAAAATGGAGGTT 1102
Qy 725 TTAAGCTTGGAGAGCTTTTGTACTATGTGGTGAATGTGTTTCAAGCCATAATCTGGCAAG 784
Db 1103 TTCGAAAGGATATGTTGCTTTATGTTATGACTTTTGGTTTGGACTTCAATAGCCTGGCAGG 1162
Qy 785 CATTTTTTGTGGAGCTATTGGGTTGATCTTCTGTGCAATGCTCTGTGCTCTCTGGAATTA 844
Db 1163 TATGCTCTGTGGTGTGGTGTGCTTGATCTTCTCTAGTGTCTTCTCTACTCCAAATGTTA 1222
Qy 845 TGGTCAGTGTCTGTCTCCGGTGAAGGTGATCTTTGGCCGTCAATTTGCTTTCAGAGAGAAT 904
Db 1223 TAAGCACAGTTTCTTTTAGCCGTAACTCTCTATTGCTGTGTATAGTTTTCATGATAAGA 1282
Qy 905 TTCAGCCGGGAAAGGTGTGGTCTCTCTCCCTCTGGGGATCAGTCTCTTATTTCT 964
Db 1283 TGAATGGGGTGAAGATAATTTCTATGCTTTTGGCTCTATGGGGTTTTCCTCTCTATATTT 1342
Qy 965 AT 966
Db 1343 AT 1344

RESULT 8

US-10-437-963-82061/c
; Sequence 82061, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 82061
; LENGTH: 1513
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81525C.1
US-10-437-963-82061

Query Match 14.4%; Score 151; DB 17; Length 1513;
Best Local Similarity 48.0%; Pred. No. 5.3e-37;
Matches 430; Conservative 0; Mismatches 465; Indels 0; Gaps 0;
Qy 71 CTCCTAATCATGCTCTTACTTCCAAATGTGGGAAAGGATCTGGTTTCCAAAGCTTCC 130
Db 1200 CACTTCTTGGAGGTTTACTACATCAAGCGGCAATGAGGATGTCACACTCG 1141
Qy 131 TTCAAACGTTGGTTGTCCACTCAATTTTCTTCCCTCTTCTCTATCTTTCTCCCGGTC 190
Db 1140 TCCAAACTGTGGCTTCCGATTTTGTTCATTGCGCTATTCTTTTCCATTCAAAGACAT 1081
Qy 191 GTCGTGTGCTTGAAGAACAGAAAGCACTCAATTTTCCCTCATGAAACCTCTCTCTTTA 250
Db 1080 CTTCTACACAACACAGTCACTAGTAGTCTCGCCCTACAAATTTCTATCCCCAAAATTA 1021
Qy 251 TCGCGCTATCGTTGCTGTGGTGTGCTGCGGATTTGACAATTTACTCTACTCTTACGGGT 310
Db 1020 TGATATATGTTGCTCGGCCCTCATCATGTCTGACAGCACTTGAATGATTTCTATGGCC 961
Qy 311 TAGCTTATATCCCTGTTTCTACTCGGCTCTTGTGATCATCTCCGCGCAATTAGGCTTCACTG 370
Db 960 TACTATATCTTCCGGTCTCAACATATTCGCTCATCTGTGCTAGTCAGCTTGCCTTCAATG 901

Qy 371 CTCTCTTTGCATTTTATGGTGAAGCAAAAGTTTCACACCTTTTCACTATATAACGCTATCG 430
Db 900 CTGTCCTTCTCATATTTTCTCAATGCTCAAAAATTCACCCCTCTGATTTTCAATTCCTAG 841
Qy 431 TTTTGCTCTCACTGGTGGTGGCTAGTCTTCCCTTAACTCTGATAGTGACAAGCTTGCAG 490
Db 840 TCCTCTTACGTTTCTGCTTCACTCTTGGAGTTGATGAAGATTTCTCAGGAACTACTA 781
Qy 491 ACAGACACACAAGGAATATGTTTGGGTTTCATCATGACTCTTGGTGCAGCTCTTCTCT 550
Db 780 GTATATCACATGGGAAGTACATTTTGGGTTTCTGTTGACACTAGGGGCATCAGTACAT 721
Qy 551 ATGGGTTTATATGCCACCTTGTGAGCTTTCTTACAGAAATCTGGTCAGGAATCAGT 610
Db 720 ACTCGCTCATTTCTCTCCCTGATGCAAGTCACATTTTGAGAAGGTTATTTAAGAGGGAGACCT 661
Qy 611 ATAGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTTGCTGCGCACCTTGTCTGCGCTCG 670
Db 660 TCTCAGTTGTGTAACATGCAGATATATACAGCTCTCGTGGCAACATTTGGCTTCTCTTG 601
Qy 671 TGGGATGCTAGCGCTGGCGATTTTCAAGGTGATAGCAGAGAGCAAGAGATTTTAAAGC 730
Db 600 TTGGGTTATTTGCAAGTGGTGAATGGATGACTTTTACAAGGAGAGATGCATGCATTCAGT 541
Qy 731 TTGGAGAGTCTTTGTACTATGTGGTGTGATCTTCTGTGCATCGTCTCTGGTCTCTGGAATATGTC 790
Db 540 CTGGGAAGCTGTCTATATGTAATGACACTGCTGTGGACGGCTATATCTTGGCAGGTAGCAT 481
Qy 791 TTGTGGAGCTATTGGGTTGATCTTCTGTGCATCGTCTCTGGTCTCTGGAATATGTC 850
Db 480 CAGTTGGAGTGGTGGGATTCATCTTGTGTGTGTCATCGCTGTTTCAATGTGATAAGCA 421
Qy 851 GTGCTGTGCTTCCGGTGACCGGTGATCTTGGCGGTCAATTTGCTTCCAGGAGAGTTTCAGG 910
Db 420 CCCTAGCTCTACCCCATCATTCCTGTTTGTGCTGTGATTTTCTTTCATGACAAGATGGATG 361
Qy 911 CGGGAAAGGTGTCGCTTTCGCTCTCCCTCTGGGATCAGTCTCTTATTTCTA 965
Db 360 GAGTAAAGATTATAGCTATGCTGATGCCCATTTTGGGGAATTTATGTCATATGGCCA 306

RESULT 9

US-10-425-114-13687
; Sequence 13687, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13687
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-024-D8_FLI
US-10-425-114-13687

Query Match 13.8%; Score 144.4; DB 16; Length 1296;
Best Local Similarity 52.4%; Pred. No. 6.3e-35;
Matches 372; Conservative 0; Mismatches 326; Indels 12; Gaps 2;
Qy 4 ATGAAGATGAAGACAGTCTTGTGAATCAATAAAGTATATTTTGGCCATTTGGAACATGT 63

Db 70 AAGACGATGCGCGCCCTCTCTGCTGGGCGACCAACTGCGTGATGCTCGCGCTTGGCGGTACCC 129
Qy 64 GGAGGCGCTCTAATGATGCGCTCTCTACTTCCAAAATGGTGGCGAAGAGATCTGGTTTCCA 123
Db 130 GGTGGGCGAGTCTCTCAGCGCCCTCTACTTTAGCAAGGCGGACACCGGAGTGCGTCTCC 189
Qy 124 AGCTTCCTTCAAACCGTTGGTGTGCACTCATTTTCTTCCCTCTCTCTTATCTTTCCTC 183
Db 190 GSAATGGCTCCAGACCGGTGGCTGGCGGCTGCTCTCCCGCAGTGGCGGCTTCTTACGTC 249
Qy 184 CGCGCTGCTGTTGCTTGAAGAACAAAGAACACCTCCATTTTCTCTCATGAAACCTCCT 243
Db 250 CGCGCGGAGCGCGC---GACCGAGCGCCCGGTGCTGCTCACCCAGACCCAGCGCGC 306
Qy 244 CTCTTATCGCCGCTATGTTGTGTTGCTCGTGGGATTTGACAAATTAACCTCTACTCT 303
Db 307 ATACTGCTGGCGCGCGGGGCTCGGCTCATCGCGGGCTGGAACAATTTGCTCTAAGCC 366
Qy 304 TAGCGGTAGCTTATATCCCTGTTTCTACTGCGTCTTTGATGATCTTCGGGCAATTAGGC 363
Db 367 TGGGGCTCTCAGTTCCTGCGCCGCTCCACCTCCCGCATCCTCATCTCCAGCAGCTGGCC 426
Qy 364 TTCACCTGCTCTCTTTGCAATTTTATGGTGAACAAAGTTTACACCTTTTCACTATAAAC 423
Db 427 TTACCGTGTCTGTTGCGGTACCTGATCGTCCGCGAGCGGCTGACGATGGCGAGCGTGAAC 486
Qy 424 GCTATGCTTTGCTCACTGCTGGTGGCGGTAGTCTTTGCCCTTAACTCTGATAGTGACAAG 483
Db 487 GCGGTGGCGCTGCTGACGGTGGCGCGCTGCTGCGGCTGCGATGCTCTCTCTGACCGC 546
Qy 484 CTTTGAAAACAGACACAAAGGAATATGTTGTTGGGTTCATCATGATCTCTTGGTGAGCT 543
Db 547 CCGCGCGGCTCACACAGGAGTCACTGCTGGGCTTCACTGCTCACTTGGCGCGCGC 606
Qy 544 CTCTCTATGGGTTATATGCACTTGTGAGCTTCTTACAGAAATC-----T 594
Db 607 GTCTGTACGGGCTGTTCTGCGCTGGCTGAGCTCACCTACAAAGTGGCGCGCGGGGCG 666
Qy 595 GGTGACGGAATACGATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTGTGCGC 654
Db 667 GGACGCGCGGTGACGTACGCGCTAGTGGTGGAGTTGAGCTGTTGATGGGTTGCTCGCC 726
Qy 655 ACTTGTGCTGCTCGTGGGATGCTAGCGGCTGGCGGATTTCAAGGTGAT 704
Db 727 ACCGCGTTCTGACCGCTCGCATGATCGTCAACAAAGGACTTTCAGGTGAT 776

RESULT 10

US-10-437-963-91425
; Sequence 91425, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 91425
; LENGTH: 1177
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_90000C.1
US-10-437-963-91425

Query Match 13.5%; Score 142; DB 17; Length 1177;
Best Local Similarity 48.4%; Pred. No. 3.5e-34;
Matches 469; Conservative 0; Mismatches 485; Indels 16; Gaps 2;
Qy 22 CTTGTAATCATAACTGATATATCTTTGGCCATTTGGAAACTGTGGAGCCCTCTTAATGATG 81
Db 142 CTGCTGGTATATTCAGCGCGTGCCTCTCTCTCGGCGCGCGCGCGCTCTCTCTC 201
Qy 82 CGTCTCTACTTCCAAAATGGTGGGAAAGATCTGTTTCCAAAGCTTCTTCAAAACGGTT 141
Db 202 CGGCTCTACTTCTGTGACCGCGGGACCCGGCTGTGCTGTCCGCCACGCTCCAGATCTCC 261
Qy 142 GGTGTGCACTCATTTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 201
Db 262 GGTGTGCGCTCTGCTGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 314
Qy 202 GAAGAAACAAGAAACGACTCCATTTTCTCATGAAACCTCTCTCTTATTCGCGCGTATC 261
Db 315 -----GCACGGATCGGCAACCTCTCTCCGCGCGCGCTCTCTCGCGCGCGCGC 366
Qy 262 GTTGTGTTGCTCTGTTGGATTTGACAAATTAACCTCTACTCTTACGGGTTAGCTTATATC 321
Db 367 GTGCTCGCGGCGTGTACCGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 426
Qy 322 CTTGTTTCTACTGCTCTTTTGATCATCTCCGCGCAATTAGGCTTCACTGCTCTCTTTTGA 381
Db 427 CCGCTGTCCACGCTGCTGCTGTGCGGACGCGAGCTGGCTTCAACCGCGGTGTTTCGCG 486
Qy 382 TTTTATGTTGGAAGAAAGTTTCACTTTTCACTATAAAACGCTATCTGTTTGTCTACT 441
Db 487 TTTCTCTTGTGGGCTCGGTTTCACTGCGGCTTCTCGGCAACGCGCTCTGCTGCTCACT 546
Qy 442 GGTGTGCGGTAGTCTTCCCTTAACTCTGATAGTGAACGCTTGAACACGAGACACAC 501
Db 547 GTGAGGAGCGGTGTCGAGGAAGCGGTGTAAGACGCTGCGGGAACCGCATACCGGTGA 606
Qy 502 AAGGAATATGTTGTTGGGTTTCATCATCTTGTGTGAGCTCTTCTCTATGGGTTTATA 561
Db 607 GGTGTGCACTTAGTGCAGCTCCAGCAAGTGGTCACTCTACAGTCCGAGATTTGGGATAG 666
Qy 562 TTGCGCACTTGTGCGAGCTTTTCAAGAAATCTGTCAGCGAAATCACTATACGTCGCG 621
Db 667 ATGAGTGTGGCGGCTTCTAGGTGGAATTTGCGGCTTACGAGCCCTTACGCGACGCTG 726
Qy 622 CTCGAGTTCAGATGCTTATGCTTGTGCTGCTTGT -GTCTGCTCTGCGGATGCT 680
Db 727 ATGAGATGACGCGGTGATGGCGCGGACACCGCGGTGTGCTGCTCGGATGCG 786
Qy 681 AGCGCTGCGGATTTCAAGGTGATAGCAGGAGAAAGCAAGAGATTTTAAGCTTGGAGAGTC 740
Db 787 GATCAAGGCGCTTTCCAGCGGCTGCGCGGGAAGCGGCGGTTCGGGCTCGCGCGCGC 846
Qy 741 TTTGTACTATGTTGATTTGTTTCAAGCAATCTGCGCAAGATTTTGTGTTGGAGC 800
Db 847 CAACTACTACCTGCTCTGCGCTGCGGACCGCTGTGCTGCGAGCTGCTCAACCTGGCAT 906
Qy 801 TATTGGGTTGATCTTCTGTGATCGCTCTGCTCTGGAATATGCTCAGTCTGCTGCT 860
Db 907 CATGGGCTCATCACTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
Qy 861 TCCGCTGAGCGGTGATTTGGCGCTCATTTGCTTCCAGGAGAAAGTTTCAAGCGGGAAGG 920
Db 967 GCGGCTCTCGCAGGCTCTCGCGCTCATCTTCTTCCAGGAAGTTTTCGACGGGACGAGG 1026
Qy 921 TGTGCTTGTGCTCTCTCCCTCTGCGGATCAGTCTCTTATTTCTATGACAGGTTAAATC 980
Db 1027 CATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1086
Qy 981 CGAGGAGAG 990
Db 1087 GAAGAAGAG 1096

```
RESULT 11
US-10-416-898-9
; Sequence 9, Application US/10416898
; Publication No. US20040172670A1
; GENERAL INFORMATION:
; APPLICANT: Walker, Elsiebeth
; APPLICANT: Dellaporta, Stephen
; TITLE OF INVENTION: MAIZE YELLOW STRIPEL AND RELATED GENES
; FILE REFERENCE: 44574-5106-US
; CURRENT APPLICATION NUMBER: US/10/416,898
; CURRENT FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: PCT/US01/43101
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 60/249,222
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 83698
; TYPE: DNA
; ORGANISM: Zea mays
US-10-416-898-9

Query Match      12.7%; Score 133; DB 17; Length 83698;
Best Local Similarity 48.0%; Pred. No. 3e-30;
Matches 455; Conservative 0; Mismatches 475; Indels 17; Gaps 2;

Qy 20 TTCTGTGTAATCAATAAAGTGTATATCTTGGCGATGGAAGAACTGTGGAGCCCTCTTAATGA 79
Db      |||
Qy 80 TCGGCTCTACTCCAAATGTCGCGAAGATGCTTCCAGCTTCCTTCAAACCG 139
Db      |||
Qy 67295 GTCGGTCTTATACAAATGAAGTGGAAACAGTAATGATCTACTCTTGTTCAAACTG 67354
Db      |||
Qy 140 TTGG-TTGTCCACTCATTTCTCCCTCTCTCTATCTTCTCCGCGGTCGCTGC 198
Db      |||
Qy 67355 GTGGCTTTCCCATCTTATCTCCCTCTTCTCTCCCTCTGCTTCAACATCTCTCTT 67414
Db      |||
Qy 199 CTTGAAGAACAGAAACGACTCCATTTTTCCTCATGAACCTCTCTTATCGCCGCT 258
Db      |||
Qy 67415 CTT-----CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 67458
Db      |||
Qy 259 ATCGTTGTTGTTGCTCGTGGATTTCACAAATACCTCTACTCTTACGGGTAGCTTAT 318
Db      |||
Qy 67459 CTTTCTCTGTTGTTGCTATTGTTAGATTAATTTTATATCTCTGTTGGACTTTTGAT 67518
Db      |||
Qy 319 ATCCCTGTTTCTACTCGCTCTTTTGATCATCTCCGCGCAATTAGGCTTCACTGCTCTTT 378
Db      |||
Qy 67519 CTCTCTGCTTCAACTTATTCGATTCATATGCTTACAGTTAGCTTTCAATGGTCTTTC 67578
Db      |||
Qy 379 GCAATTTTATGGTGAAGCAAAAGTTACACTTTCATATAAAGCTATCGTTTGTCTC 438
Db      |||
Qy 67579 TATTATTACATCAATCTCAGAAATACATTTGTTGATTTTCTTCTAGTGTGTTCTC 67638
Db      |||
Qy 439 ACTGGTGTGCGGTAGTCTTGGCCCTTAACCTCTGATAGTACAGCTTGGAAACGAGACA 498
Db      |||
Qy 67639 TCTATCTCTGCTGTTGGTTTCTTTCAGCATGATTCAAATAGCCCATCAGGAGATTCT 67698
Db      |||
Qy 499 CACAAGGAATATGTTGTTGGGTTTCATCATGACTCTTGGTGCAAGCTCTTCTCTATGGGTTT 558
Db      |||
Qy 67699 AAGTGGAGTACTTGATTTGGGTTTCTGTCAGTTTGTCTCTTATCTATCTCTT 67758
Db      |||
Qy 559 ATATTGCCACTTGTGCGAGTTTCTTAAAGAAATCTGGTCAAGCAATCACGTATACGCTC 618
Db      |||
Qy 67759 CAGCTCTCTCTATGACGTTTCTTTCGAGAAGGTTCTCAAGAGTGAAGACTCTCTCTATG 67818
Db      |||
Qy 619 GCGCTCGAGTTCACAGATGCTTATGCTTTCGCTGCACTTCTGCTGCTCTGCTGGGATG 678
Db      |||
Qy 67819 GTTCTCGAGATGCMAATCTATACGTCGCTTGTGGCTTCTTGTGTAGCGGTTATCGGATTG 67878
Db      |||

RESULT 12
US-10-437-963-72452
; Sequence 72452, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 72452
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72830C.1
US-10-437-963-72452

Query Match      12.4%; Score 130.4; DB 17; Length 789;
Best Local Similarity 53.1%; Pred. No. 1.4e-30;
Matches 306; Conservative 0; Mismatches 261; Indels 9; Gaps 1;

Qy 20 TTCTTGTAAATCAATAAAGTGTATATCTTGGCCATTGGAAACTGTGGAGGCCCTCTAATGA 79
Db      |||
Qy 218 TCCTCGTGGGCTCAACTCGCGGATGCTGGTGGGCACACGCGCGGCGCGATCCTTA 277
Db      |||
Qy 80 TGCTCTCTACTTCCAAATGTCGGAAGGATCTGTTTCCAAAGCTTCCTTCAAACCG 139
Db      |||
Qy 278 GCCGGCTCTACTTTCAGCAAGGCGGCCACAGGAAGTGGCTCTCGGCGTGGCTCGAGACCG 337
Db      |||
Qy 140 TTGGTTGTCACATTTTCTTCCCTCTCTTCTTATCTTCTCCGCGCTCGTCTGTTGCC 199
Db      |||
Qy 338 CCGGGTGGCGCTGCTGCTGTCGCGCTGTCCTACACAGCCGCGCGCGCGCG 397
Db      |||
Qy 200 TTGAGAACAGAAACGACTCCATTTTCTCATGAACCTCTCTCTCTTTATCGCCGCTA 259
Db      |||
Qy 398 ACCGCGCG-----CGCCGCTGTTCTCTACGCGCGCGCGCTGCTCTGCGGCG 448
Db      |||
Qy 260 TCGTGTGTTGCTCGTGGGATTTGACAAATACCTCTACTCTTACGGGTTAGCTTATA 319
Db      |||
Qy 449 CGGTGCTCGGGTCTTCCCGCGCGCGAGCACTTCTGCTACGCTACGGCTCGCCTACC 508
Db      |||
Qy 320 TCCCTGTTTCTACTGCGTCTTTGATCATCTCCGCGCAATTAGGCTTCACTCTCTTTG 379
Db      |||
```

Db 509 TGGCGGTGCCACCTCGGCATCTCTCCACGAGCTCGGTTACCGTCTTCTTCG 568
Qy 380 CATTTTATGTTGAAGCAAAAGTTACACACCTTTCACTATAAAGCGTATCGTTTGTCTCA 439
Db 569 CGTGCCCTCATCGTGGCGAGCGGCTACCGCGGCGAGCGTGAACCGCGTGGCGTGTCTCA 628
Qy 440 CTGGTGGTGGCGTGTGCTTGGCCCTTAACCTGTATAGTGAACAGCTTGGCAACGAGACAC 499
Db 629 CTATCGGCGCGTGTGCTGGGCTCCACGCGTCCAAAGGACCGACCCGCGGGGTGACCA 688
Qy 500 ACAAGGAATATGTTGTTGGGTTCAATCATGACTCTTGGTGCAGCTCTTCTCTATGGTTTA 559
Db 689 CGGGAAGTACTGATGGGCTCTTCTCACCTCGCGCGCGCGCTGTATCGGCTCA 748
Qy 560 TATTGCCACTGTGCGAGCTTTCTTACAAGAAATCTG 595
Db 749 TCTGCGGCTCGTGGAGCTCGGTACAGACGCCG 784

RESULT 13
US-10-424-599-31364
; Sequence 31364, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 31364
; LENGTH: 1373
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_128327C.1
US-10-424-599-31364

Query Match 12.4%; Score 130.2; DB 16; Length 1373;
Best Local Similarity 48.5%; Pred. No. 2.3e-30;
Matches 360; Conservative 0; Mismatches 383; Indels 0; Gaps 0;
Qy 258 TATCGTTGTTGTTGCTCGTGGATTTGACAAATTAACCTCTACTCTTACGGGTAGCTTA 317
Db 500 TATTGTTTGGGTTTCTTAAGTGTCTGTGATAACCTCATGTATGCTTATGCTATGCTTA 559
Qy 318 TATCCCTGTTTCTACTGCGTCTTTGATCATCTCCGCGCAATTAGGCTTCACTGCTCTTT 377
Db 560 CCTCCTGATCACCGCTCTACTGCGCTTCAATCCCTTGTGTTTCTGCACTCTT 619
Qy 378 TGCATTTTATGTTGAAGCAAAAGTTACACCTTTCACTATAAAGCGTATCGTTTGTCT 437
Db 620 TGGACTATTTCTGTGAAAAACAAAGTGAATGCTTCGATAGTAATTCGTTTTCATCAT 679
Qy 438 CACTGGTGGCGGTAGTCTTGGCCCTTAACCTGTGATAGTGAACAGCTTGGCAACGAGAC 497
Db 680 AACCAATTGCAATGACCATCATTTGCACTGGACTCGAGTTTCAGACAGATATGCCAATCAG 739
Qy 498 ACACAGGAATATGTTGTTGGTTCATCATGACTCTTGGTGCAGCTCTTCTATGGGTT 557
Db 740 TGACAGTGAATACATCGGGAATTTGTTGGGATGTTTATGATCTGCTCTTACGGGCT 799
Qy 558 TATATTGCCACTGTGCGAGCTTTCTTACAAGAAATCTGGTCAGCGAATCAGCTATACGCT 617
Db 800 TATATTGCTCTCTCGGAGCTTGTCTTGTGAAGTGTCTTGAAGAGATCTTCTATCGT 859
Qy 618 CGCGCTCGAGTTCCAGATGGTCTTATGCTTTCGCTGCCACTTGTGCTCGTGGGAT 677

Db 860 TGTCTGAGCAGCAAGTCATGGTTTCTCTGTTTGCATTTCTGTTTACCACTGTAGGGAT 919
Qy 678 GCTAGCCGCTGCGGATTTCAAGTGTATAGCAGGAGCAAGCAAGATTTTAAAGCTTGGAGA 737
Db 920 GATTATGATGTGATTTTCCAGGGATGGCAGATGAGGCTACCACTTTCAAAGGTGGTAG 979
Qy 738 GTCTTTGTACTATGTGGTGAATTTGTTTCAAGGCAATTAATCTGGCAAGCAATTTTGTGG 797
Db 980 AAGTCTTATTATCTTGTATCATTTGGGTGCAATCACTTTCACTGGGGTTCGGG 1039
Qy 798 AGCTATTCGGTGTGATCTTCTGTGCATCGTCTCTGCTCTCTGCAATTAATGTCAGTCTCT 857
Db 1040 GGGCACTGCTGTAATTTTCTTGGGCTCTACTGTGTAGCAGGTGCTTAATGCAAGTAA 1099
Qy 858 GCTTCGGGTGAGCGTGAATCTTGGCGCTCATTTGCTTCCAGGAGAAAGTTTTCAGGCGGAA 917
Db 1100 AACCCCATACAGCATTGCAGCTGTATTAAGGACCCCTATGATGGTGTTCNA 1159
Qy 918 AGGTGCTCTTGGCTCTCTCCCTCTGGGATCAGTCTCTTATTTCTATGGACAGTTAA 977
Db 1160 GATCCTCTCCCTAGTGATCACCTTTTGGGATTTGGCTCATATATTTATGGAAGTTCTAA 1219
Qy 978 ATCCGAGGAGAACTAAGGCTC 1000
Db 1220 AATATGGCCATAGACTCCAGTTTC 1242

RESULT 14
US-10-425-114-10574
; Sequence 10574, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 10574
; LENGTH: 1333
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700944981_FLI
US-10-425-114-10574

Query Match 12.1%; Score 127; DB 16; Length 1333;
Best Local Similarity 48.7%; Pred. No. 2.4e-29;
Matches 346; Conservative 0; Mismatches 365; Indels 0; Gaps 0;
Qy 258 TATCGTCTGTTGCTCGTGGATTTGACAAATTAACCTCTACTCTTACGGGTAGCTTA 317
Db 462 TATCGTTTGGGTTTCTTAAGTGTGCTGATGAACCTCATGTATGCTTATGCCATGCTTA 521
Qy 318 TATCCCTGTTTCTACTGCGTCTTTGATCATCTCCGCGCAATTAGGCTTCACTGCTCTCT 377
Db 522 CCTCCTGCATCCACTGCTCACTGTGGCATCATCCCTTGTGTTTTCGGCGCTCTT 581
Qy 378 TGCATTTTATGTTGAAGCAAAAGTTTCAACCTTTTCACTATAAAGCGTATCGTTTGTCT 437
Db 582 TGGACTATTTCTTGTGAAGAAACAAAGTGAATGCTTCAATAGTGAATTCGTTTTCGTCAT 641
Qy 438 CACTGGTGGTGGCGTAGTCTTGGCCCTTAACCTCTGATAGTGAAGCTTGGCAACGAGAC 497
Db 642 AACCGCTGCATTTGACCATCATTTGCCCTGGACTCGAGTTTCAGACAGATATCCAGCATTAG 701
Qy 498 ACACAGGAATATGTTGTTGGGTTTCATCATGACTCTTGGTGCAGCTCTTCTATGGGTT 557

```
Db 702 TGACAGTGAATACATCATGGGATTTGATGGATGTTTAGGATCTGCTTCCATGGGCT 761
Qy 558 TATATTGCCACTTGTGAGCTTTCTTACAAGAAATCTGGTCAGCGAATCAAGTATACGCT 617
Db 762 TATTTTCGCTCTCTCAGAGCTCGTCTTTGTGAAGTTGCTCGGAAGAAGATCCTTTATCGT 821
Qy 618 CCGCTCGAGTTCCAGATGGTCTTATGCTTTGCTGCCACTTGTGCTGCCCTGGGGAT 677
Db 822 TGTTCTGGAGCAGCAAGTCATGGTTTCTATTTCGATTTCTGTTTACCATCTAGGGAT 881
Qy 678 GCTAGCCGCTGGCGATTTCAAGGTGATAGCAGGAGAAGCAAGAGATTTTAAGCTTGGAGA 737
Db 882 GATTCTGAGTGTGATTTCAAGGGATGACATGAGGCTACCATCTTTCGAAGTGGTAG 941
Qy 738 GTCTTTGACTATGTGGTGAATGTGTTCAAGCCCAATATCTGGCAAGCAATTTTGTGGG 797
Db 942 AAGTCTTATTATCTTGTATCATTTGGGGTGCATCACTTTTCAGCTGGGGTTCTGGG 1001
Qy 798 AGCTATTGGGTTGATCTTCTGTGCATCTCTCTGCTCTCTGCTCTGCAATTATGTCAGTCTCT 857
Db 1002 GGGCACTGCTATAATTTCTTGGGCTCTACTGTGCTAGCAGTGTGCTTAATGCAGTAAG 1061
Qy 858 GCTTCCGGTGAAGGTGATCTTGGCGTCAITTTGCTTCCAGGAGAAGTTTTCAGGCGGGAA 917
Db 1062 AACACCCATAACAAGCATTTGCAGCTGTTTACTGCTAAAGGACCCCTATGAGTGTTCAA 1121
Qy 918 AGGTGTCGCTTTGGCTCTCTCCCTCTGGGATCAGTCTCTTATTCTATGG 968
Db 1122 GATCCTCTCCCTAGTAGTATCACCTTTTGGGGATTTGGCTCATATATTTATGG 1172
```

RESULT 15

```
US-10-437-963-43860/c
; Sequence 43860, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 43860
; LENGTH: 2381
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2381)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_46977C.1
US-10-437-963-43860
```

```
Query Match 11.3%; Score 118.8; DB 17; Length 2381;
Best Local Similarity 47.6%; Pred. No. 1.4e-26;
Matches 351; Conservative 0; Mismatches 387; Indels 0; Gaps 0;

Qy 259 ATCGTTGTGGTGTGCTCGTGGGATTTGACAAATTAACCTTACTCTTACGGGTTAGCTTAT 318
Db 2021 ATCGTCTGGGCTCATCATCGCCGCGACGACATGATGTACACTGTGTGGCTCAAGTAC 1962
Qy 319 ATCCCTGTTTCTACTGGGTCTTTGATCATCTCCGCGCAATTAAGCTTTCACCTGCTCTCTT 378
```

```
Db 1961 CTCCCCGCTTCGACCTACTCGCTCATCTGTGCCAGCAGCTCGCGTTCAATGTGCTCTTC 1902
Qy 379 GCATTTTTTATGGTGAAGCAAAAGTTTCAACCTTTTCACTATAAAACGCTATCGTTTTGCTC 438
Db 1901 TCATACGTGTCTCAACTCCAGAGGTCACTCTCTGATATTTCAACTCCGTGCTGCTGCTC 1842
Qy 439 ACTGGTGGTGGCGTAGTCTTCCCTTAACTCTCATAGTAGTACAAAGCTTGGAAAACGAGACA 498
Db 1841 ACCATGTCCGCTTCGCTCATCGGAGTCAGCAAGAGTCTCAGGGGGTACCGGGCTCTCG 1782
Qy 499 CACAAGGAATATGTTGTTGGGTTTCATGACTCTTGGTGCAGCTCTTCTCTATAGGGTTT 558
Db 1781 GGAGGGAAGTATCTGCTCGGTTTCGTGCTGACGCTGGGGGCGTCTGCACCTACTCTCGCTG 1722
Qy 559 ATATTGCCACTTGTGAGCTTTCTTACAAGAAATCTCTGGTCAGCGAATCAAGTATACGCTC 618
Db 1721 ATCTCTCGGCTGATGACAGCTCACCTTCGAGACCATCATCAAGAGACACACCTTCTCAGCC 1662
Qy 619 GCGCTCGAGTTCACAGATGGTCTTATGCTTTGCTGCCACTTGTGCTCGCTCGTGGGATG 678
Db 1661 GTCTCAACATGACAGATCTACACGGCGCTCGTGGCGACGGCCGCTCGGTGGTGGGCTG 1602
Qy 679 CTAGCCGCTGGCGATTTCAAGGTGATAGCAGGAGAAGCAAGAGATTTTAAAGCTTGGAGAG 738
Db 1601 TTCCGAGCGCGAGTGGAGTCTGCTGAGGGGGGAGATGAACGCTTCAGTCCGGGCAG 1542
Qy 739 TCTTTGTACTATGTGTGATTTGTTCACGGCCCAATATCTGGCAAGCAATTTTTTGTGGGA 798
Db 1541 TTCTCTACCTGATGACGCTGCTGTGGGCGGCGGTGTCGTGGCAGGTGGCCAAACATCGGG 1482
Qy 799 GCTATTGGGTTGATCTTCTGTGATCGTCTCTGCTCTCTGAAATTATGGTCAGTCTCTG 858
Db 1481 GTGCTCGGCTCATCTTCGAGGTGTCGGCGCTCTTCTCCAACGTGATCAGCACGGTGTCA 1422
Qy 859 CTTCGGGTGACGGTGAATCTTTGGCGCTCATTTGCTTCCAGGAGAAGTTTTCAGGGCGGAAA 918
Db 1421 CTGGCGTCTATCCGCTTCTTCGCGGTGGTGTGTTCCACGACAGGATGAACGGGTGAAG 1362
Qy 919 GGTGTGCTGCTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTATTTCTATGAGCAGGTTAAA 978
Db 1361 ATCGTGGCCATGCTGATTCGAATTTGGGGATTTTATTCGTATCTGTTTTCAGCACTATCTA 1302
Qy 979 TCCGAGGAGAAGACTAAG 996
Db 1301 GATGGCAAGAAAGCGAAG 1284
```

Search completed: November 1, 2004, 23:39:51
Job time : 553.317 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:45:29 ; Search time 3737.17 Seconds
(without alignments)
10228.408 Million cell updates/sec

Title: US-09-913-767-2
Perfect score: 1049
Sequence: 1 aaagatgaagatgaagacagt.....actgattatgtagcttaaaa 1049

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	999.8	95.3	1230	3	CNS0A99D
2	521	49.7	827	8	BZ515292 Arabidops
3	515	49.1	527	8	BZ515292 Arabidops
4	417.2	39.8	745	8	AQ958362 LERAX12TR
5	414.4	39.5	1059	5	AQ958361 LERAX12TR
6	410.8	39.2	688	8	BH589732 BGS0504TR
7	334.4	31.9	463	5	BX836413 BX836413
8	308.2	29.4	985	7	CK257466 EST741103
9	306.2	29.2	986	7	CK243751 EST727388
10	304.6	29.0	680	4	BG439880 GA_Ea000
11	302.2	28.8	998	7	CK247918 EST731555
12	301.4	28.7	1013	7	CK251896 EST735533
13	297.2	28.3	907	7	CK257887 EST741524
14	290.6	27.7	979	7	CK252005 EST735642
15	288.4	27.5	988	7	CK248243 EST731890
16	285.2	27.2	971	7	CK272858 EST718936
17	281.8	26.9	645	1	AU238506 AU238506
18	280.4	26.7	930	7	CK253429 EST737066
19	278.8	26.6	929	7	CK249872 EST733509
20	277.4	26.4	942	7	CK249994 EST733631
21	274.6	26.2	908	7	CK247701 EST731338
22	272.4	26.0	932	7	CK256828 EST740465
23	268	25.5	948	7	CK257752 EST741389
24	266.4	25.4	899	7	CK256093 EST739730

25	266	25.4	928	7	CK245031	CK245031
26	265.2	25.3	688	1	AJ807511	AJ807511
27	262.6	25.0	894	7	CK247243	CK247243
28	261.8	25.0	868	7	CK245030	CK245030
29	258.6	24.7	882	7	CK258020	CK258020
30	258.6	24.7	889	7	CK251126	CK251126
31	256.8	24.5	880	7	CK257437	CK257437
32	255	24.3	892	7	CK246384	CK246384
33	250.8	23.9	784	8	BH601144	BH601144
34	249.6	23.8	722	4	BM436674	BM436674
35	242.8	23.1	359	8	BH441465	BH441465
36	239	22.8	832	7	CK259060	CK259060
37	236	22.5	703	6	CB006438	CB006438
38	233	22.2	829	7	CK262517	CK262517
39	226.4	21.6	827	7	CK249190	CK249190
40	222	21.2	902	6	CK262518	CK262518
41	220	21.0	665	6	CD714562	CD714562
42	219.8	21.0	588	5	BU762991	BU762991
43	215.2	20.5	616	5	BQ873555	BQ873555
44	215.2	20.5	701	5	BQ874152	BQ874152
45	215	20.5	668	6	CB008408	CB008408

ALIGNMENTS

CNS0A99D 1230 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTSL372D07 of Silique of strain col-0 of Arabidopsis thaliana
(thale cress).
ACCESSION BX821376.1 GI:42468600
VERSION HTC; GSLT_cDNA.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1230)
AUTHORS Castell, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1230)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full-length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES
source

1. 1230
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"

QY 301 TCCTACGGGTAGCTTATATCCCTGTTTCTACATGCGTCTTTTGATCATCTCCGCGCAATTA 360
|||||
Db 504 GCTTACGGGTAGCTTACCTCCAGATTTCACATCTCTTTGGTCTATCTCTCTCCCAACTG 445
QY 361 GGCTTCACATGCTCTCTTTGATTTTATGTTGTTGAAGCAAAAGTTTCAACCTTTCACATA 420
|||||
Db 444 GCCTTCACATGCTCTCTTTGCTTTTTCATGGTGAAGCAAAAGTTTCAAGCAATTTATTATA 385
QY 421 AACGCTATCGTTTGTCTCACTGGTGGTGGCGTATGCTTCCCTTAACTCTGATAGTGAC 480
|||||
Db 384 AACGCCGTGTTTGTCTTACTCTGGGTGGCGGAATCTTTGCCCTTCAACCGAGGTGAC 325
QY 481 AAGCTTGCAAAACGAGACACACAAGGAATATGTTTGGGTTCATCATGACTCTTTGGTGCA 540
|||||
Db 324 AAGCTTGCAAAACGAGACACACAAGGAGTATATCGTTGGGTTCCTCATGACCCCTTGGTGG 265
QY 541 GCTCTTCTCTATGGGTTTATATGCCACTTGTGAGCTTCTTCAAGAANAATCTGTGAC 600
|||||
Db 264 GCTATTCTCTACCGCTTGTATTGGCGCTTGTGGAGCTCACTTCAAGAAAGCTCGTCAA 205
QY 601 CGAATCAGGTATACGCTCGCGCTCGAGTTCAGATGCTTCCAGATGCTTATGCTTGTGCGCACTGT 660
|||||
Db 204 CGAATCAGTACAGCTTGTGCTCGAGATGCACTTGGTCTTGTGCTTGTGCGCACTTGC 145
QY 661 GTCTGCTCGTGGGATGCTAGCCGCTGGCGATTTCAAGGTGATAGCAGGAGCAAGA 720
|||||
Db 144 TTCTGCTCGTGGGATGCTAGCTGATGGCGATTTCAAGGAGTT-----AGCAAGA 94
QY 721 GATTTTAACTTGGAGCT-----TTGTACTATGTTGGTATGTTGTTCAAGGCAATA 774
Db 93 GATTTTAACTTGGAGGCTTACTACGTATTACTATGTTGGTGGTGTGTTTCAAGGCAATC 34
QY 775 ATCTGCAAGCATTTTGTGGGAGCTATTGGG 807
Db 33 ATCTGCAAGGCTTTTCTTGGATCTATTGGG 1

RESULT 3
AQ958362/c
LOCUS
DEFINITION
LERA Arabidopsis thaliana genomic clone LERAX12, genomic
survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 527)
Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utterbach, T.,
Feldblyum, T., Liang, F., Creasy, T. and Fraser, C.M.
Genomic survey sequencing of Landsberg erecta ecotype of
Arabidopsis thaliana and identification of sequence-based
polymorphisms
Unpublished (2000)
Contact: Xiaoying Lin
The Institute for Genomic Research
7912 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: at@tigr.org
For additional information, see <http://www.tigr.org/tdb/at/at.html>
Seq primer: TR
Class: shotgun.

FEATURES
source
1..527
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Landsberg erecta"
/db_xref="taxon:3702"

/clone="LERAX12"
/clone_lib="LERA"
/note="Organ: Leaf; Vector: phos1; Total genomic DNA was
sheared to 0.9-1 kbp before ligation."

ORIGIN

Query Match 49.1%; Score 515; DB 8; Length 527;
Best Local Similarity 99.0%; Pred. No. 6.4e-143;
Matches 518; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 91 TTCCAAATATGGTGGCGAAGGATCTGGTTTCCAAAGCTTCTCAACACCGTTGGTTGTCCA 150
Db 523 TCCAAAAATGGTGGCGAAGGATCTGGTTTCCAAAGCTTCTCAACACCGTTGGTTGTCCA 464
QY 151 CTCATTCTTCCCTCTCTCTTATCTTTCCTCCGCGCGTCTGTTGCTTCAAGAACAA 210
Db 463 CTCATTCTTCCCTCTCTTATCTTTCCTCCGCGCGTCTGTTGCTTCAAGAACAA 404
QY 211 GAAACGACTCCATTTTCTCATGAAACCTCTCTCTTTATCGCGCTATCGTTGTGGT 270
Db 403 GAAACGACTCCATTTTCTCATGAAACCTCTCTCTTTATCGCGCTATCGTTGTGGT 344
QY 271 TTGCTCGTGGGATTTGACAAATTACCTTACTCTTAGGGGTAGCTTATATCCCTGTTCT 330
Db 343 TTGCTCGTGGGATTTGACAAATTACCTTACTCTTAGGGGTAGCTTATATCCCTGTTCT 284
QY 331 ACTGCGTCTTTGATCATCTCCGCGCAATTAGGCTTCTACTGCTCTCTTTGCAATTTTATG 390
Db 283 ACTGCGTCTTTGATCATCTCCGCGCAATTAGGCTTCTACTGCTCTCTTTGCAATTTTATG 224
QY 391 GTGAAGCAAAAGTTTCAACCTTCTACTATAAAGCTATCGTTTGTCTCACTGGTGGTGC 450
Db 223 GTGAAGCAAAAGTTTCAACCTTCTACTATAAAGCTATCGTTTGTCTCACTGGTGGTGC 164
QY 451 GTAGTCTTGGCTTAACTCTGATAGTACAGCTTGAAGAGACACACAGGAATAT 510
Db 163 GTAGTCTTGGCTTAACTCTGATAGTACAGCTTGAAGAGACACACAGGAATAT 104
QY 511 GTTGTGGTTCATCATCTTGGTGCAGCTCTTCTCTATGGGTTTATATATGACACTT 570
Db 103 GTTGTGGTTCATCATCTTGGTGCAGCTCTTCTCTATGGAATATATATGACACTT 44
QY 571 GTGAGCTTTTCTTACAAGAAATCTGGTCAGCGAATCAGGTATA 613
Db 43 GTGAGCTTTTCTTACAAGAAATCTGGTCAGCGAATCAGGTATA 1

RESULT 4
AQ958361
LOCUS

DEFINITION
LERA Arabidopsis thaliana genomic clone LERAX12, genomic
survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 745)
Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utterbach, T.,
Feldblyum, T., Liang, F., Creasy, T. and Fraser, C.M.
Genomic survey sequencing of Landsberg erecta ecotype of
Arabidopsis thaliana and identification of sequence-based
polymorphisms
Unpublished (2000)
Contact: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: at@tigr.org

For additional information, see <http://www.tigr.org/tdb/at/at.html>
Seq primer: TR
Class: shotgun.

FEATURES
source
1..527
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Landsberg erecta"
/db_xref="taxon:3702"

For additional information, see <http://www.tigr.org/tdb/at/at.html>

Seq primer: TF
Class: shotgun.

FEATURES

source
1. 745
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="landsberg erecta"
/db_xref="taxon:3702"
/clone="LERAX12"
/clone_lib="LERA"
/note="Organ: Leaf; Vector: pHOS1; Total genomic DNA was sheared to 0.9-1 Kbp before ligation."

ORIGIN

Query Match 39.8%; Score 417.2; DB 8; Length 745;
Best Local Similarity 99.1%; Pred. No. 1.7e-113;
Matches 419; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 AAGATGAAGATCAAGACAGTCTTGTATCATAACTGTATATCTTGGCCATTTGGAAC 60
Db |||||
323 AAGATGAAGATGAAGACAGTCTTGTATCATAACTGTATATCTTGGCCATTTGGAAC 382
Qy 61 TGTGGAGCCCTCTAATGATCGCTCTCTACTTCCAAATGTTGGCGAAAGGATCTGGTT 120
Db |||||
383 TGTGGAGCCCTCTAATGATCGCTCTCTACTTCCAAATGTTGGCGAAAGGATCTGGTT 442
Qy 121 CCAAGCTCTCTCAAAACGGTGTGTCACATATTTCTCCCTCTCTCTATCTTTC 180
Db |||||
443 CCAAGCTCTCTCAAAACGGTGTGTCACATATTTCTCCCTCTCTCTATCTTTC 502
Qy 181 CTCGCCGCTCGTGTGCTTGAAGACAAAGACGATCAATTTCTCTCATGAACCT 240
Db |||||
503 CTCGCCGCTCGTGTGCTTGAAGACAAAGACGATCAATTTCTCTCATGAACCT 562
Qy 241 CCTCTCTTTATCGCGCTATCGTGTGTTGCTCGTGGGATTTGACAAATTAACCTCTAC 300
Db |||||
563 CCTCTCTTTATCGCGCTATCGTGTGTTGCTCGTGGGATTTGACAAATTAACCTCTAC 622
Qy 301 TCTTACGGGTTAGCTTATATCCCTGTGTTCTACTCGCTTTTGATCATCTCCGCAATTA 360
Db |||||
623 TCTTACGGGTTAGCTTATATCCCTGTGTTCTACTCGCTTTGATCATCTCCGCAATTA 682
Qy 361 GCCTTCACTGCTCTTTGCAATTTTATGTTGGAAGAAAGTTTCACACCTTTCACATA 420
Db |||||
683 GCCTTCACTGCTCTTTGCAATTTTATGTTGGAAGCAAAAGTTACACCTTTTCACTATA 742
Qy 421 AAC 423
Db |||
743 AAC 745

RESULT 5
BX841065
LOCUS
DEFINITION BX841065 Arabidopsis thaliana Flowers and buds Col-0 Arabidopsis thaliana cDNA clone G51F65ZD02 SPRIM, mRNA sequence.
ACCESSION BX841065
VERSION BX841065.1 GI:42535148
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1059)
Castelli, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, P., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
Unpublished (2004)

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences).
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES

Location/Qualifiers
source
1. 1059
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="GSLTFB65ZD02"
/tissue types="Flowers and buds"
/clone_lib="Arabidopsis thaliana Flowers and buds Col-0"

ORIGIN

Query Match 39.5%; Score 414.4; DB 5; Length 1059;
Best Local Similarity 65.7%; Pred. No. 1.3e-112;
Matches 661; Conservative 0; Mismatches 321; Indels 24; Gaps 3;
Qy 1 AAGATGAAGATGAAGACAGTCTTGTATCATAACTGTATATCTTGGCCATTTGGAAC 60
Db |||||
27 AAGAAGAAGATGAAGAATGGTTTGTATATCATAACTGTATATCTTGGCCATTTGGAAC 86
Qy 61 TGTGGAGCCCTCTAATGATCGTCTCTACTTCCAAATGTTGGCGAAAGGATCTGGTTT 120
Db |||||
87 TGTGGAGTCTCTTGTGTTAACTCGTCTCTACTTCCAAATGTTGGCGAAAGGATCTGGTT 146
Qy 121 CCAAGCTCTCTCAAAACGGTGTGTTGTCACATCTTCTTCTCTCTCTCTCTCTCTCTCT 180
Db |||||
147 ATGAGCTTCTCATCAACCGTGGTTTTTCCAATCATCTCATCTCTCTCTCTCTCTCTCT 206
Qy 181 CTCGCCGCTCGTGGTTG-----CCTTGAAGACAAAGAACGACTC 220
Db |||||
207 CTCAGCCGTAGCGCGAGCAACCGCAACCCCTAACACCGCGAATAAAGCGGAACAA 266
Qy 221 CATTTTTCCTCATGAACCTCCTCTCTTATCGCGCTATCGTTGTTGTTGTTGTTGTTGTT 280
Db |||||
267 GCTACTTCTCATGGAAACCTCTCTGTTATCGCTCCATTTGTCATAGGGTTGCTCACAG 326
Qy 281 GATTGACAAATTAACCTCTACTCTTAACGGGTTAGTTATATCCCTGTTTCTACTGCTCTT 340
Db |||||
327 GACTTGACAACTACTTATATTTCTTACGGATAGCATATCTGCCAGTTTCAACTTCATCGC 386
Qy 341 TGATCATCTCCGCGCAATTAGGCTTCACTGCTCTCTTTCATTTTATGTTGTTGTTGTTGTT 400
Db |||||
387 TCATAATCGGAACCTCAACTAGCTTTTCAACGCTCTCTTCGCTTCTTGAATTAATCAAGCAA 446
Qy 401 AGTTTCAACCTTTTCACTATATAAAGCTAT-CGTTTGTCTCACTGTTGTTGTTGTTGTTGTT 459
Db |||||
447 AGTTCACTCGTTTCTCCATAAACCGCGTACGTTTGTGACGGTTGTTATCGGGATCCTT 506
Qy 460 GCCCTTAACCTCTGATAGTGAACAGCTTGGAAACAGACACAGGAATATGTTGTTGGG 519
Db |||||
507 TCTTTACACAGTAATGAAGAAAAAACCCTCTTAAGAAAAAATAAATAATAATGTTGTTAA 566
Qy 520 TTCAATCATGACTCTTGGTGAGCTCTTCTCTATGGTTTATATGTTGTTGTTGTTGTTGTT 579
Db |||||
567 TTCTCGAATGACTTTTTTTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 626
Qy 580 TCTTACAGAAATCTGTTGACGGAATCAAGTATACGCTCGCTCGAGTTCCAGATGGTC 639
Db |||||

Db 627 AATTAAAAAAGATCGTCAAGAAATCACCTTTCCCACTGTGCTTGAGATTCAGATGGTC 686
Qy 640 TTATGCTTTGCTGCCACCTTGCTGCTCGTGGGATGCTAGCCGCTGCGCATTTCAAG 699
Db 687 ATGCGCTTCTTCTAATTTTCTTTGTCATTTTCATTTTCATCTTTGGAGATTTAAG 746
Qy 700 GTGATAGCAGGAGCAAGAGATTTTAACTTTGGAGAGTCTTTG---TACTATGGTGG 756
Db 747 GTGAAGCAAGCAAGCAAGATGTTCAAGATTTGGTGGATCATTTGTTTACTATTCAATT 806
Qy 757 ATTGCTGTTACGGCCATAATCTGGCAAGCATTTTGTGGAGCTATTGGGTTGATCTTC 816
Db 807 TTGTGATCACAGATATATGCGCAAGTTTCTTTAGTGTCCAAAGGGAATGTGTTT 866
Qy 817 TGTGATCGTCTCTGCTCTCGAATATATGGTCAGTCTCTGCTCCGGTGCAGCGTGATC 876
Db 867 TGTGATCATCACTAGCTTCTGCTGTTCTGATAAGTGTCTGCTCCGGTGACTGAAGTT 926
Qy 877 TTGCGCGTCATTTGCTTCAGGAGAAGTTTCAGCGGGGAAAGGTGTCGCTTTGGCTCTC 936
Db 927 TTGCGCGTCGTTGTTTTCGGTGTGATGTTTCAGTTAGTGTAAAGGTGCTCTACTTCTT 986
Qy 937 TCCTCTCGGGATCAGTCTCTTATTTCTATGGACAGTTAAATCCG 982
Db 987 TCACTTTGGGTTTGTCTCATACATCTAAGGTGTGTTTAAATCGG 1032

RESULT 6

BH589732/c

LOCUS BH589732 688 bp DNA linear GSS 15-DEC-2001
DEFINITION BOGES04TR BOGE Brassica oleracea genomic clone BOGES04, genomic survey sequence.

ACCESSION BH589732

VERSION BH589732.1 GI:17842184

KEYWORDS GSS

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 688)

Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished (2001)

Other GSSs: BOGES04TF

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

FEATURES

source

Location/Qualifiers
1..688
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO100DH3"
/db_xref="taxon:3712"
/clone="BOGES04"
/db_lib="BOGE"
/note="Vector: pHOS1; Site 1: BatX1; 2-3 kb sheared genomic DNA inserted into pHOS1 using BatX1 linkers"

ORIGIN

Query Match 39.2%; Score 410.8; DB 8; Length 688;
Best Local Similarity 77.3%; Pred. No. 1.4e-111;
Matches 512; Conservative 0; Mismatches 147; Indels 3; Gaps 1;
Qy 1 AAGATGAAGATGAAGACATCTTGTATATATAAATCTGTATATCTTGGCCATTTGAAC 60
Db 664 ATGATGAAGATGAAGCGGCTTCTTGTATATATAAATCTGTATATATATTTGGCGATAGAAAC 605

Qy 61 TGTGGAGGCCCTCTAATGATGGTCTCTACTTTCCAAATGTTGGCGAAAGATCTGTTT 120
Db 604 TGTGGAGGCCCTCTAATGATGGTCTCTACTTTCCAAATGTTGGCGAAAGATCTGTTT 545
Qy 121 CCAAGTCTCTTCAAAACGGTGTGTTGTCACATTTTCTCCCTCTCTCTATATCTTC 180
Db 544 TCAAGCTTCTCTCAAAACCTCAGGTTGCCCATCATCTCTCCCTCTCTCTCTCTTC 485
Qy 181 CTCGCCGCTGCTGTTGCTTGAAGCAAGAAAC---GACTCCATTTTCTCATGAA 237
Db 484 CTCACCGGTACCGTAAACAAGACAGAAATACGAAGACTACACTCTTCTCATCAA 425
Qy 238 CTTCTCTCTTTATCGCGCTATCTGTTGTTGTTGCTCGTGGGATTTGACAAATACCTC 297
Db 424 CTTCTCTGTTCTCGCTCTGTTGTTGTTGCTCTCTCATCGGCTTCGACAACTACCTC 365
Qy 298 TACTCTTACGGGTTAGCTTATATCCCTGTTTCTACTGCGTCTTTGATCATCTCCGGCAA 357
Db 364 TACGCTTACGGGTTAGCTTATACCTCCCGTCTCCAGCTCTCTCTCATCATCTCTCTCAG 305
Qy 358 TTAGGCTTCACGCTCTCTTTGCATTTTATGTTGGAAGCAAAAGTTTACACCTTCACT 417
Db 304 CTAGCCTTACCGCTCTCTTTCGCCCTTCTTCATGGTGAAGCAAAAGTTTACGCTTTC 245
Qy 418 ATAAACGCTATCGTTTGTCTCACTGCTGCTGCGCTAGTCTTTCGCCCTTAACTCTGATAGT 477
Db 244 ATAAACGCGCTGTTTGTCTCACTCTCGGCGCGGAGCCTTGTCTTTCACGGGACGT 185
Qy 478 GACAAGCTTGAAAACAGACACACAGAAATATGTTGTTGGGTTTCATCATGACTTTGTT 537
Db 184 GACAAGCTTCCCAAGGAGACGCGAAGAGTATAGTTGGGTTTGTATGCGGTCGCT 125
Qy 538 CGAGCTCTCTATGGGTTTATATGTCACCTTGTGCGAGCTTCTTACAGAAATCTGTT 597
Db 124 CGAGCTGTCTCTACGCTTGTGTTGCGCTTGTGGAGCTTACTTACAAAGAAAGTACT 65
Qy 598 CAGCGAAATCAGTATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTCTGCGCACT 657
Db 64 CAGAGATCAGCTACACGCTTGTGGTGTGATGATGCAAGTTGGTCTTGTGCGTGTGCGCACT 5
Qy 658 TG 659
Db 4 TG 3

RESULT 7

BX836413/c

LOCUS

DEFINITION

BX836413 Arabidopsis thaliana Silique Col-0 Arabidopsis thaliana

cDNA clone GSLTSL532D06 3PRIM, mRNA sequence.

ACCESSION BX836413

VERSION BX836413.1

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 463)

Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,

Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,

Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:

A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished (2004)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out

full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Sallanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences).
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/EST
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES
source
Location/Qualifiers
1..463
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL53ZD06"
/tissue_type="Silique"
/clone_lib="Arabidopsis thaliana Silique Col-0"

ORIGIN

Query Match 31.9%; Score 334.4; DB 5; Length 463;
Best Local Similarity 97.4%; Pred. NO. 1e-88;
Matches 372; Conservative 0; Mismatches 6; Indels 4; Gaps 3;

Qy 672 GGGGATGCTAGCCGCTGGCGATTCAAGGTG-ATAGCAGGAGAGCAGAGATTTTAAGC 730
Db 463 GGGGATGCTAGCCGCTGGCGATTCAAGGTGATAGCAGGAGAGCAGAGATTTTAAGC 404
Qy 731 TTGAGAGCTTTTGTAATGTGGTGATTTGTTTCAAGCCATATCTGGCAAGCATTTT 790
Db 403 TTGAGAGCTTTTGTAATGTGGTGATTTGTTTCAAGCCATATCTGGCAAGCATTTT 344
Qy 791 TTGCGGAGCTATTGGGTGTGATCTTCTGTGCATCGTCTCGGTCTCTGGAAATTAAGTCA 850
Db 343 TTGCGGAGCTATTGGGTGTGATCTTCTGTGCATCGTCTCGGTCTCTGGAAATTAAGTCA 284
Qy 851 GTGCTCTGCTCCGCTGACGCTGATCTT--GGCGTCAATTTGCTTCCAGGAGAGTTTCA 908
Db 283 GTGCTCTGCTCCGCTGACGCTGATCTTCCGCGTGCAATTTGCTTCCAGGAGAGTTTCA 224
Qy 909 GCGGGGAAAGGTGTGCTTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTATTTCTATGG 968
Db 223 GCGGGGAAAGGTGTGCTTTGGCTCTCTCCATCTGGGGATCAGTCTCTTATTTCTATGG 164
Qy 969 ACAGGTTAAATCCGAGGAGAGACTAAGGCTCAGATACACAATGCTCAGCTTCCAGT 1028
Db 163 ACAGGTTAAATCCGAGGAGAGACTAAGGCTCAGATACACAATGCTCAGCTTCCAGT 104
Qy 1029 T-ACTGATTATAGCTTAAAA 1049
Db 103 TCACGTAAATGTAGCTTAAAA 82

RESULT 8

CK257466
LOCUS
DEFINITION
EST741103 potato callus cDNA library, linear EST 30-JUL-2004
Solanum tuberosum cDNA clone POCD557 5' end, mRNA sequence.

ACCESSION

CK257466

VERSION

CK257466.1 GI:39814446

KEYWORDS

EST.

SOURCE

Solanum tuberosum (potato)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 985)
Buell.C.R., Hart.A., Zismann.V., Karamycheva,S.A. and Baker,B.
Generation of ESTs from potato callus tissue
Unpublished (2003)
Contact: Robin Buell
The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via <http://genome.arizona.edu/orders/>.
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

Location/Qualifiers
1..985
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCD557"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 29.4%; Score 308.2; DB 7; Length 985;
Best Local Similarity 58.4%; Pred. NO. 9.9e-81;
Matches 560; Conservative 0; Mismatches 393; Indels 6; Gaps 1;

Qy 3 GATGAAGATGAAGACAGTTCTTCTTAATCATAACTGTATATCTTGGCCATTGGAAACTG 62
Db 33 GAAATGAAGAAGCTTATCTTCTTAATTTTCAACATAAATATGCTTTCAATCGGCAATTG 92
Qy 63 TGGAGGCCCTCTAATGATGGTCTCTACTTCCAAAATGGTGGGAAAGAGTCTGGTTTCC 122
Db 93 CGGTGGTCCCTTAATCTCTCGACTCTATTTCATGCGGGCAACGAATTTGGATACC 152
Qy 123 AAGTTCCTTCAAAACCGTTGGTGTCCACTCAATTTTCTCCCTCTTCTCTTAICTTTCT 182
Db 153 CAGCGTATTACAAACTGTTGGTTGTCCAATTTTCTCATCCCTCTAGCCATAGCTATTT 212
Qy 183 CCGCCGTCGTCGTTCCTTGAAGAACAAGAACGACTCCATTTTCTCATGAACCTCC 242
Db 213 CCAACGCCGGAATTC-----AAGACCGGAGGCTAAATCGTCTTATACACGGCA 266
Qy 243 TCTCTTTATCCCGCTATCGTTGTGTTTGTCTCGTGGGATTTGACAAATACCTCTACTC 302
Db 267 AGAGTTCAATCGCATCTGCGGCGTGGGAATCATCGTCGGTCTCGATGTTTAAACTC 326
Qy 303 TTAGGGTTAGCTTATATCCCTGTTTCTACTGGGTCTTTGATCATCTCCGGCAATTAGG 362
Db 327 ATGGGGACCGCGAAATFACCCGTTTCAACTTCGACATAATCAACCGGACTCAACTGC 386
Qy 363 CTTCACTGCTCTCTTTGCAATTTTATGGTGAAGCAAAAGTTCCACACCTTTTCATATAAA 422
Db 387 GTTCACTGGCGCTTTTGTCTGTCTTATAGTTAAACAGAAATTTGACAGGTATTCGACGA 446
Qy 423 CGCTATCGTTTGTCTCACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 482
Db 447 TTCCGTGTTTGTCTTATCCCGGAGCTGCGACTTTAGCTCTCCGGGCGAAGCGTGACCG 506
Qy 483 GCTTGCAAAACGACACACAAAGGAATATGTTGTTGGTTTCATCATGATCTCTGGTGAGC 542
Db 507 GCGGCGCGCGAGTCCACGAAGGATTATATGTTGGGGTTTGTGATGACGTTTATCGGTGC 566
Qy 543 TCTTCTCTATGGGTTTATATTGCCACTTGTCCAGCTTTCTTACAGAAATCTGCTCAGCG 602
Db 567 GGTGTTATAGGATTGATTGTTGCCGTTAATGTAGTTGATTATATGAAGCGGAAGCAAGC 626
Qy 603 AATCACGTATACGCTCCGCTCGAGTTCCAGATGGTCTTATGCTTGTCTGCCACTTGTGT 662
Db 627 TGTACTTACACTACAGTATTGGAGATTTCAGATGGTTTGGGCAATTTCTCTACTGTTT 686
Qy 663 CTGCCTCGTGGGATGCTAGCCGTGGCGATTTTCAAGGTGATAGCAGGAGCAAGCAAGA 722
Db 687 TTGCACCTATTGGAATGGTTTATCAACAGGATTTTCCAGCGGATTTCCAGGGAAGCAAGTCA 746


```
Db 319 ATGGGACCGCGAAATACCCGTTTCACTTCGACATAATCAACGCGACTCAACTGC 378
QY 363 CTTCACTGCTCTCTTGGCATTTTATGGTGAAGCAAAAGTTTCACACCTTTTCACTATAA 422
Db 379 GTTCACTGCGCTTTTGTGCTGCTTATAGTAAACAGAAATTCAGACGTAATTCGACGAA 438
QY 423 CGCTATGCTTTTCTCACTGCTGCTGCGGAGTTCCTTGGCCCTTAACTCTGATAGTACAA 482
Db 439 TTCGTCGTTTTCCTTATCGCCGAGCTGCGACTTTAGCTCTCCGGCGAAGCGTGACCG 498
QY 483 GCTTGCACAGCAGACACACAGGAATATGTTGGGTTTCATCATCACTCTTGGTCGACG 542
Db 499 GCCGCGCGGAGTCACAGGAATATATGTTGGGGTTTGTGATCAGCTTATCGGTCG 558
QY 543 TCTTCTCTATGGGTTTATATGTCACCTGTCGAGCTTTCTTACAAGAAATCTGGTCAGCG 602
Db 559 GGTGTTATATGAGTATGATGTTGCGGTTAATGAGTTGATTTATATGAGCGGAGCAAGC 618
QY 603 AATCACGTATACCTCGCGCTCGAGTTCAGATGCTTATGCTTTTGTGCCACTTGTGT 662
Db 619 TGTACTTACACTACAGTATTCAGATTCAGATGTTTGGGCAATTTCTGCTACTGTTT 678
QY 663 CTGCTCTGCGGAGTGTAGCCGCTCGCGATTTCAAGGTGATAGAGGAGCAAGAGA 722
Db 679 TTGCACATTTGGAATGGTTATCAACAAGGATTTCCAGGCGATTTCCAGGGAAGCAAGTCA 738
QY 723 TTTTAAAGCTTGGAGAGTCTTTGTTACTATGTTGTTGATGTTGTTCCAGCGCATATCGGCA 782
Db 739 ATATGCAATGGAGAACTTAAGTATTTATATGTTAGTATGTTGTTGCTATTTTGGCA 798
QY 783 AGCAATTTTGTGGGAGCTATTGGGTTGATCTTCTGTGCATCGTCTCTGCTCTGGAAT 842
Db 799 ATTCCGCTCTCGGGTGTAGTGGAGTTATTTTATTTCTTCTCTCTCTCTGGAAT 858
QY 843 TATGTCAGTGTCTCTGCTCCGTTGACGCTGATCTTTGGCGCTCAITTTGCTTCAGAGAA 902
Db 859 TATAGCGCTTTTATCTCTGTTACCGAAGTTTGGGCTGTAATTTGTTCCAGCAAAA 918
QY 903 GTTTCAGGCGGGAAGGTGTCCTTTGGCTCTCTCCCTCTGGG 947
Db 919 ATTTCAAGCGGAAAAGGAGTAGCTATTTTCTAGCTATATGGG 963
```

RESULT 12

CK251896 1013 bp mRNA linear EST 30-JUL-2004
EST35533 potato callus cDNA library, normalized and full-length
Solanum tuberosum cDNA clone POCBR03 5' end, mRNA sequence.

CK251896

CK251896.1 GI:39805365

EST.

Solanum tuberosum (potato)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.

1 (bases 1 to 1013)

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

Generation of ESTs from potato callus tissue

Unpublished (2003)

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via <http://genome.arizona.edu/orders/>

Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

1..1013

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Kennebec"

/db_xref="taxon:4113"
/clone="POCBR03"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 28.7%; Score 301.4; DB 7; Length 1013;
Best Local Similarity 58.2%; Pred. No. 1.1e-78;
Matches 569; Conservative 0; Mismatches 401; Indels 7; Gaps 2;

QY 3 GATGAAGATGACAGAGTTCTTGTAAATCAATAAATGTTATCTTGGCCCATTTGGAACATG 62
Db 40 GAAATGAAGAAGCTTATCTTCTTAATTTTCAACATAATATATGCTTTCAATCGGCAATG 99

QY 63 TGGAGGCCCTCTAATGATGCTCTCTACTTCCAAATGTTGGGCAAGGATCTGGTTTCC 122
Db 100 CGTGGTCTTTAATCTCTGACTCTATTTTCAATTCATGCGCGCCAAAGAAATTTGGATACC 159

QY 123 AAGCTTCTTCAAACCGTTGGTTGCTCACTCAATTTTCTTCCCTCTTCTCTTATCTTTCT 182
Db 160 CAGCGTATTTACAAACTGTTGGTTGTCCAATTTATCTCATCTCCCTCTAGCCATAGCTTAT 219

QY 183 CGCGCTGCTGCTTGGCTTGAAGACAGAAAGCACTCCATTTTCTCTCATGAACTCC 242
Db 220 CCAACGCCGGAATAATTC-----AAGGACCGAGGCTAAATCGTCTTTATCACAGGCA 273

QY 243 TCTCTTTATCGCGCTATCTGTTTGTGTTTGTCTGTTGGGATTTGACAAATTAACCTCTAC 302
Db 274 AGAGTTTCATCGCATCTGCGCGGCTCGGAATCATCTGCTCGCTCTCGATGTTTACTTAA 333

QY 303 TTACGGTTAGCTTATATCCCTGTTTCTACTCGCTCTTTTGATCATCTCCGCGGAATAGG 362
Db 334 ATGGGGACCCCGGAAATTAACCGTTTCAACTTCGACACTTAATCAACGCGACTCAACTGC 393

QY 363 CTTCACTGCTCTCTTTTGGCAATTTTATGTTGAAGCAAAAGTTTCAACCTTTCACTATAA 422
Db 394 GTTCACTGCGCTTTTGTGCTTATAGTTAAACAGAAATTGACAGCGTATTCGACGAA 453

QY 423 CGCTATCGTTTGTCTCACTGTGTGCGCTAGTCTCTGCTTAACTCTGATAGTACAA 482
Db 454 TTCCGCTGTTTGTCTTATCGCGGAGCTGCGACTTTAGCTCTCCGGGCAACGTTGACCG 513

QY 483 GCTTCAAAACGAGACACACAGGAATATGTTGTTGGGTTTCACTGACTCTTGGTGACAG 542
Db 514 GCGCGCCGCGAGTCCACGAGGATATATGTTGGGTTTGTGATGACGCTTTATCGGTGC 573

QY 543 TCTTCTATGGGTTTATATGCGCACTTGTGCGAGCTTCTTACAGAAATCTGGTCAGCG 602
Db 574 GGTGTTATATGGAATGATGTTTGGCGTTAAATGAGTTGATTTATATGAAGCGAAGCA 633

QY 603 AATCAGGTATAGCTCGCGCTCGAGTTCAGATGCTTATGCTTTGCTTGTCCACTTGTGT 662
Db 634 TGTACTTACACTACAGTATTCGAGATTCAGATGTTTGGGCAATTCGCTACTGTTT 693

QY 663 CTGCTCTGTTGGGATGCTAGCGCTGGCGATTTTCAAGGTGATAGCAGGAAAGCAAGAGA 722
Db 694 TTGCACATTTGGAATGGTTATCAACAGGAATTTCCAGGCGATTTCCAAGGGAAGCAAGTCA 753

QY 723 TTTTAAAGCTTGAGAGTCTTTGTACTATCTGTTGTTGTTTTCACGCCCATATCTGGCA 782
Db 754 ATATGCAATTTGAGAAATCTAAGTATATATGTTAGTATGTTGGTGTGCTTATTTGGCA 813

QY 783 AGCAATTTTGTGGGAGCTATTGGGTTGATCTTCTGTGTCATCTCTCTGCTCTCTGGAAT 842
Db 814 ATTCGCGCTCCGGGGTAGTTGGAGTTATTTTATTTCTTCTCTTACTCTCTGGAAT 873

QY 843 TATGTCAGTGTCTGCTTCCGTTCCGTTGATCTTGGCCGTCATTTGCTTCCA-GGAGA 901


```

source
1. .979
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POC8R86"
./tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cdna library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

```

ORIGIN	Query Match Best Local Similarity Matches	27.7%; 58.2%; 551; Conservative	Score 290.6; Pred. No. 1.9e-75; 0; Mismatches 389;	DB 7; Length 979; Indels 7; Gaps 2
Qy	3	GATGAAGATGAAGACAGATCTTGTAAATCATAAACGTATATCTTGGCCATTGGAAACTG	62	
Db	39	GAATAATGAAGAGCTTATCTTCCCTAAATTTCAACATAAATATATGCTTTCAATCGGCAATTG	98	
Qy	63	TGGAGGCCCTCTAATGATGCGTCTCTACTTCCAAATGCTGGCGAAAGATCTGTGTTCC	122	
Db	99	CGGTGGTCTTAACTCTCGACTCTATTTCATTCATGGCGGCCAAGAAATTTGGATACC	158	
Qy	123	AAGCTTCCTCAAAACGGTTGGTTGCCACTCAATTTCTTCCTCTCTCTTAATCTTTCT	182	
Db	159	CAGCGTATTACAAACGTGTGGTTGTCCAATTAATCTCATCCCTAGCCATAGCCTATTT	218	
Qy	183	CCGCGCGCTGCTGGCTTGAAGACAGAAACGACTCCATTTTCTCATGAACACTCC	242	
Db	219	CCAAACGCCGAAATTC-----AAGGACCGAGGCTAAATCGTCTTTATCACACGGCA	272	
Qy	243	TCTCTTTATCCCGCTATCGTTGTGTGGTTGCTCGTGGGATTTGCAATTAATCTCTACTC	302	
Db	273	AGATTTCATCGATCTCGCGCTCGGAATCATCGTCGTCTCGATGGTTACTTAACTC	332	
Qy	303	TTACGGGTTAGCTTATATCCCTGTTTCTACGTGCTTTTGATCATCTCGCGCAATTAGG	362	
Db	333	ATGGGGACCCGAAATATACCGTTTCAACTTCGACTCTAATCAACGGGACTCAACTTGC	392	
Qy	363	CTTCACATGCTCTCTTTGCAATTTTATCGTGAAGCAAAAGTTTCACACCTTTACACTATAA	422	
Db	393	GTTTCACTGCGCTTTTGTGCTTATAGTTTAAACAGAAATTCAGCGGTATTCGACGAA	452	
Qy	423	CGCTATCGTTTTGCTCACTGTGTGGTCCGTAGTCTCTGCCCTTAACTCTGATAGTAGCAA	482	
Db	453	TTCCGTGTTTTGCTTATCGCGGAGCTCGCACTTTTAGCTCTCGGGCGAACGGTGACCG	512	
Qy	483	GCTTGCMAACGAGACACACAGGAATATGTTGTGGGTTTCATCATGACTCTTTGGTGCAGC	542	
Db	513	GCCGCGCGGCGAGTCCACGAAGGATATATGTGTGGGGTTTGTGATGAGGTTATCGGTGC	572	
Qy	543	TCCTCTCTATGGGTTTATATGCCACTTGTTCGAGCTTTCTTACAGAAATCTGTGTCAGCG	602	
Db	573	GGTGTATATGGAATGATGTTGCCGTTAATTGAGTTGATTTATATGAAGCGCAACGAC	632	
Qy	603	AATCACGFPATACGTCGCGCTCGAGTTCAGATGGTCTTATGCTTTTGTGCGCACTTGTGT	662	
Db	633	TGTTACTTACACTACAGTATTGGAGATTTCAGATGGTTTTTGGSCATTTCTGCTACTGTTT	692	
Qy	663	CTGCGCTGTGGGGATGCTAGCCGCTGGCGATTTCAAGGTGATAGCAGGAGACGAAGAGA	722	
Db	693	TTGCACACTATGGAATGATTTACAAACAGGATTTCCAGGCGATTTCCAGGGAAGCAAGTCA	752	
Qy	723	TTTTAAGCTTCGAGAGCTTTTGTACTATGTGTGATTTGTTCACGGCCATAATCTGGCA	782	
Db	753	ATATGCAATTGGAGNAICTAGTATATATATGGTACTAGTATGGTGTGCCATTAATTGGCA	812	
Qy	783	AGCATTTTTTGTGGGAGCTATTGGGTGTGATCTTCTGTGCATCTGTCGTCTCTCGGAAT	842	

Db 813 ATTCCGCGCTCTCTGGGGTAGTTGGAGTTATTTTTATTTCTTCTCTTACTCTCTGGAAT 872

Qy 843 TATGGTCAGTGCCTCTGCTTCGGGTGACGGTGATCTTGGCCGCTCATTTTGGCTTTCACAGGAGAA 902

Db 873 TATAGGCGCTTTTTTACTTCTGCTGTTTACGAAAGTTTGGCTGTAATTT-TTTTACGAAAA 931

Qy 903 GTTTCAGCGGGGAAAGGTTCGCTTTGGCTTCTCTCCCTCTGGGGAT 949

Db 932 ATTTCAAGCGGAAAAAGAGTAGCTATTTTCTAGCTCTATGGGGAT 978

RESULT 15

CK248243 988 bp mRNA linear EST 30-JUL-2004

LOCUS EST731880 potato callus cDNA library, normalized and full-length

DEFINITION Solanum tuberosum cDNA clone POCAZ44 5' end, mRNA sequence.

ACCESSION CK248243

VERSION CK248243.1 GI:39798187

KEYWORDS EST.

SOURCE Solanum tuberosum (potato)

ORGANISM Solanum tuberosum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.

TITLE Generation of ESTs from potato callus tissue

JOURNAL Unpublished (2003)

COMMENT Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.

Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source Location/Qualifiers

1..988

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="POCAZ44"

/tissue_type="callus"

/lab_host="DH10B-Tona"

/clone_lib="potato callus cDNA library, normalized and full-length"

/note="Vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI; supplier: RNA was isolated from Solanum tuberosum var. Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 27.58; Score 288.4; DB 7; Length 988;

Best Local Similarity 57.88; Pred. No. 8.7e-75;

Matches 554; Conservative 0; Mismatches 396; Indels 8; Gaps 2;

Qy 3 GATGAAGATGAAGACAGTTCCTGTAATCATAAACTGTATATTCTTGCCCATTCGAAACTG 62

Db 39 GAAATGAAGAGCTTATCTCTCTAATTTTCAACATAATTATGCTTTCATCGCAATTG 98

Qy 63 TGAAGGCCCTCTTAATGATGCGTCTCTACTTCCAAATGTGGCGAAAGGATCTGGTTTC 122

Db 99 CGGTGGTCTTTAATCTCTCGACTTATTTCAATTCATGGCGGCAACGAATTTGGATACC 158

Qy 123 AAGCTTCCTTCAACCGTGGTTGTCCACTATTTTCCCTCTCTCTATCTTCTTCTTCCT 182

Db 159 CAGCGTATTACAAACTGTGGTTGTCCAATATTCTCATCTCCCTAGCCATAGCCTATTT 218

Qy 183 CCGCCGTCGTCGTTGCTTGAAGAAACAAGAAACGACTCCATTTTTTCTCTATGAACTCC 242

Db 219 CCAACGCCGGAATTC-----AAGGACCGAGGCTAAATCGTCTTTATCACAGGCA 272

Qy 243 TCTCTTTATCGCGCTATCGTGTGTTGGTTTGCTCGTGGGATTTGACAAATTACCTCTACTC 302

Db 273 AGAGTTTCATCGCATCTGCGGCGTCGGAAATCATCGTCGGTCTCGATGGTTACTTTAAACTC 332
Qy 303 TTACGGGTAGCTTATATCCCTGTTTCTACTGCGTCTTTGATCATCTCCGGCGAATTAGG 362
Db 333 ATGGGACCCCGGAAATTTACCCGTTTCAACTTCGACTCTAATCAACGCGACTCAACTTGC 392
Qy 363 CTTCACGTCTCTCTTTGATTTTATGGTGAAGCAAAAGTTACACACTTTTCACTATAAA 422
Db 393 GTTCACTGCGCTTTTGTGTGCTTATAGTTTAAACAGAAATTGACGGTATTCGACGAA 452
Qy 423 CGTATCGTTTGTCTCACTGCTGTGTCGCCGTAGTCTTCCCTTAACTCTGATAGTGACAA 482
Db 453 TTCGTCGTTTGTCTTATCGCGGAGCTGCGACTTTAGCTCTCCGGGGAACGGTGACCG 512
Qy 483 GCTTGCAAAACGAGACACACAAGGAATATGTTGTTGGGTTCATCATGACTCTTTGGTGCAGC 542
Db 513 GCCGCGCGCGAGTCCACGAAGGATATATGTTGGGGTTTGTGATGACGTTTATCGGTGC 572
Qy 543 TCTTCTCTATGGGTTTATATGCCACTGTGTCGAGCTTCTTACAGAAATCTGGTCAGCG 602
Db 573 GGTGTTATAGGATGATGTTGCCGTTAAITGAGTTGATTTATATGAAGCGGAAGCAAGC 632
Qy 603 AATCAGGTATACGCTCGCGCTCGAGTTCACAGATGGTCTTATGCTTTGCTGCCACTTGTGT 662
Db 633 TGTACTTACACTACAGTATTTGGAGATTCAGATGTTTGGGCATTTCTGCTACTGTTT 692
Qy 663 CTGCTCGTGGGATGCTAGCCGCTGGCGATTTCAAGGTGATGACGAGAGCAAGAGA 722
Db 693 TTGCACTATTGGAATGATTTATCAACAAGGATTTCCAGCGGATTCCAAGGGGAAGCAAGTCA 752
Qy 723 TTTTAAAGCTTGGAGAGTCTTTGTACTATGTGGTGAATGTGTTACGGGCCATTAATCTGGCA 782
Db 753 ATATGCAATTGGAGAACTAAGTATATATGGTACTAGTATGGTGTGCCATTATTGGCA 812
Qy 783 AGCATTTTGTGGGAGCTATTGGGTTGATCTTCTGTGCATCGTCTCTGGTCTCTGGAAT 842
Db 813 ATTCGCGCTCCTGGGGTAGTTGGAGTTATTTTATTTCTTCTCTTTACTCTCTGGAAT 872
Qy 843 TATGGTCAGTCTCTGCTTCCGGTGACGGTGATCTTGGCCGTCATTTTGGTTCAGGAGAA 902
Db 873 TATAGGCGCTTTTCTTCTGTTTACGAAGTTTGGCTGTAATTTTGTTCACGAAAAA 932
Qy 903 GTTTCAGCGGGGAAAGGTGTCGCTTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTAT 960
Db 933 --TTTCAGCCGAAAAGAGTAGCTATTTTCTAGCTCTATGGGATTTGTTTCATAT 988

Search completed: November 1, 2004, 22:30:03
Job time : 3748.17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:40:28 ; Search time 550.844 Seconds
(without alignments)
9996.736 Million cell updates/sec

Title: US-09-913-767-2
Perfect score: 1049
Sequence: 1 aaagatgaagatgaagacagt.....actgattatgtacttaaaa 1049

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1049	100.0	1049	3	Aaa97920 A. thalia
2	528	50.3	1225	3	Aaa97919 A. thalia
3	428.8	40.9	2586	3	Aac44184 Arabidops
4	199.2	19.0	493	3	Aac36831 Arabidops
5	167.6	16.0	3387	8	Abz14407 Arabidops
6	167.6	16.0	3387	8	Ada68246 Arabidops
7	157	15.0	1293	3	Aaa97922 A. thalia
8	155.8	14.9	1200	8	Ada70627 Rice gene
9	154.2	14.7	1071	3	Aaa97925 A. thalia
10	149	14.2	1403	3	Aac37652 Arabidops
11	148.8	14.2	1047	6	Adg88190 A. thalia
12	148.8	14.2	1081	3	Aaa97924 A. thalia
13	142.2	13.6	2660	3	Aac47974 Arabidops
14	137	13.1	1194	3	Aaa97923 A. thalia
15	133	12.7	83698	6	Abn85767 Arabidops
16	112	10.7	1145	3	Aaa97921 A. thalia
17	112	10.7	1292	3	Aac51119 Arabidops
18	110.4	10.5	1295	3	Aac37289 Arabidops
19	108.8	10.4	1152	8	Ada70213 Rice gene
20	103.2	9.8	1155	8	Ada69686 Rice gene
21	103	9.8	2175	8	Ada70748 Rice gene

22	90	8.6	1418	3	AAC46395	Aac46395 Arabidops	
23	90	8.6	1421	3	AAC39779	Aac39779 Arabidops	
24	81.6	7.8	1040	8	ADA69568	Ada69568 Rice gene	
25	74.4	7.1	592	10	ABX56692	Abx56692 Arabidops	
26	56.6	5.4	1173	12	ADN72520	Adn72520 Thale cre	
27	51	4.9	774	6	ABN99127	Abn99127 Arabidops	
28	50.2	4.8	477	3	AAC37273	Aac37273 Arabidops	
29	48.2	4.6	332	12	ADP92185	Adp92185 Cotton ex	
c	30	42	4.0	677	6	ABQ35095	Abq35095 Oligonucl
31	42	4.0	677	6	ABQ35094	Abq35094 Oligonucl	
32	40	3.8	1186	10	ADC87306	Adc87306 Human GPC	
c	33	39.2	3.7	6741	3	AAAI0595	Aaai0595 Gene enco
34	39.2	3.7	9447	6	ABL33166	Ab133166 Human imm	
c	35	38	3.6	1648	10	ADC87042	Adc87042 Human GPC
36	38	3.6	2000	8	ADA71938	Ada71938 Rice gene	
c	37	37.8	3.6	18467	4	ABL06976	Ab106976 Drosophil
38	37.4	3.6	2000	8	ADA71938	Ada71938 Rice gene	
c	39	37	3.5	110000	6	ABX08336_05	Continuation (6 of
40	37	3.5	110000	6	ABX08336_06	Continuation (7 of	
41	37	3.5	110000	12	ADJ25985_05	Continuation (6 of	
42	37	3.5	110000	12	ADJ25985_06	Continuation (7 of	
43	37	3.5	110000	12	ADN97989_05	Continuation (6 of	
44	37	3.5	110000	12	ADN97989_06	Continuation (7 of	
45	37	3.5	110000	12	ADO50281_05	Continuation (6 of	

ALIGNMENTS

RESULT 1

AAA97920

AAA97920 standard; DNA; 1049 BP.

XX

XX

AC

AAA97920;

XX

XX

DT

19-JAN-2001 (first entry)

XX

XX

DE

DE

XX

XX

KW

A. thaliana PUP1 DNA #2.

KW

PUP1; transgenic plant; nucleobase transporter; apical dominance;

XX

XX

OS

flowering behaviour; senescence; pesticide distribution; ds.

XX

XX

PN

Arabidopsis thaliana.

XX

DE19907209-A1.

XX

PD

24-AUG-2000.

XX

XX

19-FEB-1999; 99DE-01007209.

XX

XX

19-FEB-1999; 99DE-01007209.

XX

XX

(FROM/) FROMMER W.

XX

XX

Gillissen B, Buerkle L, Andre B, Frommer WB;

XX

XX

WPI; 2000-566202/53.

XX

XX

Nucleic acid, useful for producing transgenic plants with altered

PT

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

ALIGNMENTS

RESULT 1
AAA97920
ID AAA97920 standard; DNA; 1049 BP.
XX
AC AAA97920;
XX
DT 19-JAN-2001 (first entry)
XX
DE A. thaliana PUP1 DNA #2.
XX
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.
XX
OS Arabidopsis thaliana.
PN DE19907209-A1.
XX
PD 24-AUG-2000.
XX
PF 19-FEB-1999; 99DE-01007209.
XX
PR 19-FEB-1999; 99DE-01007209.
XX
(FROM/) FROMMER W.
PA Gillissen B, Buerkle L, Andre B, Frommer WB;
PI WPI; 2000-566202/53.
DR
XX
PT Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transporter, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.
PS
XX
CC Claim 2; Page 12; 24pp; German.

This invention describes a novel nucleic acid encoding a plant nucleobase transporter (I). (I) is produced by complementation of a nucleobase transport (NBT)-defective host cell with a plant gene bank by selection of NBT-positive cells. (I) is used to isolate homologous sequences from bacteria, fungi, plants, animals and humans, for expression of the encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting expression of (II) (when in antisense orientation), and to produce transgenic crop plants. The transgenic plants have modified nucleobase transport properties, e.g. altered affinity and substrate specificity

Db 16 AAGAGAGATGAAGATGTTTGGATTAATCATAACTGTTATTAATCTCACTATAGGAACA 75
Qy 61 TGTGGAGGCTCTAAATGATGCTCTACTTCCAAAATGCGTGGCGAAAGGATCTGTTT 120
Db 76 TGTGGAGGCTCTTGTAACTGCTCTACTTCCAAATGGCGGAAACGAATCTGGTTC 135
Qy 121 CCAAGTCTGCTCAAAACCGTGTGTCACACTCAATTTCTTCCCTCTCTCTTATCTTTC 180
Db 136 ATGAGCTTCTATCAACCGCTGTTTTCAAATCATCTCATCCCTCTCTTGGTCTCTTTC 195
Qy 181 CTCGCGGCTGCTGTTG-----CCTTGAAGAAACAAGACGATCCA 222
Db 196 CTCAGCGCTGCGCGGCAACCGCAACCCCTAACACCGCGGAAACAAGCGGAAACAAG 255
Qy 223 TTTTCTCTCATGAAACCTCTCTCTTATCGCGCTATCGTTGTTGTTTGTCTGTTGGA 282
Db 256 CTCCTCTCATGAAACTCTCTGTTTATCGCTCCATGTCATAGGGTGTCTACAGGA 315
Qy 283 TTTGACAAATPACTTACTCTTACGGGTAGCTTATATCCCTGTCTTCTACTCGCTCTTTC 342
Db 316 CTTGACAACTACTTATATCTTACGGATTAGCATATCTGCCAGTTTCAACTTCATCGCTC 375
Qy 343 ATCATCTCCGCGCAATTAGCTTCACTGCTCTCTTTGCAATTTTATGTTGAAGCAAAAG 402
Db 376 ATAATCGGAACCTCAACTAGCTTTCAACGCTCTCTTCGCTTCTTGTAGTCAAGCAAAAG 435
Qy 403 TTCACACCTTTCACTATAAAGCTATCGTTTCTCACTGGTGGTCCGTAGTCCCTTGGC 462
Db 436 TTCACCTCGTCTCATAAAGCGCTGTTTGTGACGGTGGTATCGGGATCTCTTGGC 495
Qy 463 CTTAACTCTGATAGTCAAGCTTGAACCGAGACACACAAGGAATATGTTGGGTTTC 522
Db 496 TTAACACAGTATGGAGACAAACCGCTAAGGAGACGAAGAGATGTTGGTGGTTC 555
Qy 523 ATCATGACTTGGTGCAGCTCTTCTCTATGGTATATATGTCACCTTGTGAGCTTTCT 582
Db 556 TTGATGACTGTGGTGCAGCTCTTCTCTATGCTTTTATATACCGCTCGTTGAGCTAACT 615
Qy 583 TACAAGAAATCTGGTCAGCGAATCAGTATAGCTCGCTCGCTGAGTTCAGATGGTCTTA 642
Db 616 TACAAGAAAGCTGTCAGAAATCACTTTCCCACTTGTGCTTGAGATTCAGATGGTCA 675
Qy 643 TGCCTTGTGCTGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 702
Db 676 TGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735
Qy 703 ATAGCAGGAGACGAAGATTTTAAGCTTGGAGATC---TTTGTACTATGTTGGTGATT 759
Db 736 ATAGCAGGAGACGAAGATTTCAAGATTTGGAGGATCAGTGTGTTTACTATGCAATTGATA 795
Qy 760 GTGTTCAAGCCCAATCTGGCAAGCATTTTGTGGGAGCTATGCGGTTGATCTTCTGT 819
Db 796 GTGATCAGAGGAATAATATGGCAAGTTCCTCTTAGGAGCCATAGGGAATGTGTTTGT 855
Qy 820 GCATCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 879
Db 856 GCATCATCACTAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 915
Qy 880 GCGCTCATTTGCTTCCAGAGAGATTTTCAAGCGGGGAAAGGTTGCTGCTTGGCTCTCTCC 939
Db 916 GCGCTGCTTGTGTTCCGGGAGAGTTTCAAGCGAGAGAAAGGTGCTCTCTACTCTCTTCT 975
Qy 940 CTCTGGGATCAGTCTCTTATTTCTATGACAGGTTAAATCC-----GAGGAG 987
Db 976 CTTTGGGATTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1035
Qy 988 AAGACTAAGGCTCAGGATACAACTGTCTCAGCTTCCAGTTTACTGATTATGTAGCTTAA 1047
Db 1036 AAACCTCAACCGCGGAGACAGACTGCTTCTTCCAGTTAGTATGTTGCTTAA 1095

AAC4184
ID AAC4184 standard; DNA; 2586 BP.
XX
AC AAC4184;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 41938.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN BP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132487P.
PR 06-MAY-1999; 99US-0132488P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.

PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
----	--------------	----------------	----	--------------	----------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db	1826	CGTTTGTGACGGTTGGTATCGGGATCCTTGGCTTACACAGTGATGGAGACAAACCGGC	1885	99US-0128234P.
Qy	489	AAACGAGACACAAGGAATATGTTGGTTCATCATGACTCTTGGTGCAGCTCTTCT	548	99US-0128714P.
Db	1886	TAAGGAGACGAAGAAGATGTTGGTGGTTCCTTGATGACTGTGGTGCAGCTCTTCT	1945	99US-0129845P.
Qy	549	CTATGGGTTTATATGCCACTTTCGAGCTTCTTACAAGAAATCTGGTCAGCGAATCAC	608	99US-0130077P.
Db	1946	CTATGCTTTTATATACCGCTCGTTGAGCTAACTTACAAGAAAGCTCGTCAAGAAATCAC	2005	99US-0130443P.
Qy	609	GTATAGCTCGCCTCGAGTTCAGATGGTCTTATGCTTGTGGTGCACCTGTGTGCGCT	668	99US-0130510P.
Db	2006	TTTCCCACTTGTGCTTGAATTCAGATGGTCACTGCTGCTGCTGCTGCTGCTGCTGCT	2065	99US-0130891P.
Qy	669	CGTGGGATGCTAGCGCTCGCCATTTTCAAGGTAGTAGGAGCAAGCAAGAGATTTTAA	728	99US-01313449P.
Db	2066	CATTGGCATGTTTCATCGTTGGAGATTTAAGGTAGTAGCAVAGAGCAAGAGATTTCAA	2125	99US-0132048P.
Qy	729	GCTTGGAGATC---TTTGTACTATGTGGTGAATTGTTTCACGGCCATTAATCTGGCAAGC	785	99US-0132407P.
Db	2126	GATTGGAGGATCAGTGTTTTACTATGATGATGATCAGAGGAATAATATGGCAAGG	2185	99US-0132484P.
Qy	786	ATTTTGTGGAGCTATTGGGTTGATCTTCTGTGCATCGTCTGCTGCTGCGAATTAT	845	99US-0132485P.
Db	2186	TTTCTTCTTAGGAGCCATAGGATTTGTTTGTGCATCATCACTAGCTTCTGCTGCTCT	2245	99US-0132486P.
Qy	846	GGTCACTGCTCTGCTCCGCTGACGGTGATCTTGGCGCTCATTTGCTTCCAGGAGATTT	905	99US-0132863P.
Db	2246	GATAAGTGTCTGCTTCCGCTGACTGAAGTTTTCGCGCTGCTGTTGTTCCGGGAGAAGTT	2305	99US-0134218P.
Qy	906	TCAGGCGGGGAAAGGTGTGCTTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTATTCTTA	965	99US-0134219P.
Db	2306	TCAGGCAGAGAAAGGTCTCTACTTCTTCTTCTTGGGATTTGCTCTTACTTCTTA	2365	99US-0134221P.
Qy	966	TGACACAGGTTAAATCC-----GAGGAGAAGACTAAGGCTCAGGATACACAACCT	1013	99US-0134221P.
Db	2366	CGCGGAGTTTAAATCCGCAAGAAAGTTGTGATAAACCTCAACCGCGGAGACAGAACT	2425	99US-0134221P.
Qy	1014	GTCTCAGCTTCCAGTTTACTGATTATGTAGCTTAA	1047	99US-0134221P.
Db	2426	GCCTATTTCCAGTTAGTATGTTGCTTAA	2459	99US-0134221P.
RESULT 4				
AAC36831				
ID	AAC36831 standard; DNA; 493 BP.			
XX	AAC36831;			
DT	17-OCT-2000 (first entry)			
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 15222.			
KW	Hybridisation assay; genetic mapping; gene expression control;			
KW	protein identification; signal transduction pathway; metabolic pathway;			
KW	promoter; termination sequence; ss.			
OS	Arabidopsis thaliana.			
XX	EF1033405-A2.			
XX	06-SEP-2000.			
XX	25-FEB-2000; 2000BP-00301439.			
XX	25-FEB-1999; 99US-0121825P.			
PR	05-MAR-1999; 99US-0123180P.			
PR	09-MAR-1999; 99US-0123548P.			
PR	23-MAR-1999; 99US-0125788P.			
PR	25-MAR-1999; 99US-0126264P.			
PR	29-MAR-1999; 99US-0126785P.			
PR	01-APR-1999; 99US-0127462P.			
PR	06-APR-1999; 99US-0128234P.			
PR	08-APR-1999; 99US-0128714P.			
PR	16-APR-1999; 99US-0129845P.			
PR	19-APR-1999; 99US-0130077P.			
PR	21-APR-1999; 99US-0130443P.			
PR	23-APR-1999; 99US-0130510P.			
PR	23-APR-1999; 99US-0130891P.			
PR	28-APR-1999; 99US-01313449P.			
PR	30-APR-1999; 99US-0132048P.			
PR	04-MAY-1999; 99US-0132407P.			
PR	05-MAY-1999; 99US-0132484P.			
PR	06-MAY-1999; 99US-0132485P.			
PR	06-MAY-1999; 99US-0132486P.			
PR	07-MAY-1999; 99US-0132487P.			
PR	11-MAY-1999; 99US-0132863P.			
PR	14-MAY-1999; 99US-0134218P.			
PR	14-MAY-1999; 99US-0134219P.			
PR	14-MAY-1999; 99US-0134221P.			
PR	18-MAY-1999; 99US-0134370P.			
PR	19-MAY-1999; 99US-0134768P.			
PR	20-MAY-1999; 99US-0134941P.			
PR	21-MAY-1999; 99US-0135124P.			
PR	21-MAY-1999; 99US-0135353P.			
PR	25-MAY-1999; 99US-0136021P.			
PR	27-MAY-1999; 99US-0136392P.			
PR	28-MAY-1999; 99US-0136782P.			
PR	01-JUN-1999; 99US-0137228P.			
PR	03-JUN-1999; 99US-0137528P.			
PR	04-JUN-1999; 99US-0137502P.			
PR	07-JUN-1999; 99US-0137724P.			
PR	08-JUN-1999; 99US-0138094P.			
PR	10-JUN-1999; 99US-0138540P.			
PR	10-JUN-1999; 99US-0138847P.			
PR	14-JUN-1999; 99US-0139119P.			
PR	16-JUN-1999; 99US-0139452P.			
PR	16-JUN-1999; 99US-0139453P.			
PR	17-JUN-1999; 99US-0139453P.			
PR	18-JUN-1999; 99US-0139454P.			
PR	18-JUN-1999; 99US-0139455P.			
PR	18-JUN-1999; 99US-0139456P.			
PR	18-JUN-1999; 99US-0139457P.			
PR	18-JUN-1999; 99US-0139458P.			
PR	18-JUN-1999; 99US-0139459P.			
PR	18-JUN-1999; 99US-0139460P.			
PR	18-JUN-1999; 99US-0139461P.			
PR	18-JUN-1999; 99US-0139462P.			
PR	18-JUN-1999; 99US-0139463P.			
PR	18-JUN-1999; 99US-0139750P.			
PR	18-JUN-1999; 99US-0139763P.			
PR	21-JUN-1999; 99US-0139817P.			
PR	22-JUN-1999; 99US-0139899P.			
PR	23-JUN-1999; 99US-0140353P.			
PR	23-JUN-1999; 99US-0140354P.			
PR	24-JUN-1999; 99US-0140695P.			
PR	28-JUN-1999; 99US-0140823P.			
PR	29-JUN-1999; 99US-0140991P.			
PR	30-JUN-1999; 99US-0141287P.			
PR	01-JUL-1999; 99US-0141842P.			
PR	02-JUL-1999; 99US-0142154P.			
PR	06-JUL-1999; 99US-0142390P.			
PR	08-JUL-1999; 99US-0142803P.			
PR	09-JUL-1999; 99US-0142920P.			
PR	12-JUL-1999; 99US-0142977P.			
PR	13-JUL-1999; 99US-0143542P.			
PR	14-JUL-1999; 99US-0143624P.			
PR	15-JUL-1999; 99US-0144005P.			
PR	16-JUL-1999; 99US-0144085P.			
PR	16-JUL-1999; 99US-0144086P.			
PR	19-JUL-1999; 99US-0144325P.			
PR	19-JUL-1999; 99US-0144331P.			

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2212.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
XX
PR 26-JAN-2001; 2001US-0264647P.
XX
PR 22-JUN-2001; 2001US-0300111P.
XX
XX (SCRI) SCRIPPS RES INST.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Harper JF, Kreps J, Wang X, Zhu T;
XX
XX WPI; 2002-304127/34.
XX
XX Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
XX Claim 144; SEQ ID NO 2212; 577bp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
XX Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;
XX
Query Match 16.0%; Score 167.6; DB 6; Length 3387;
Best Local Similarity 48.6%; Pred. No. 5.7e-42;
Matches 492; Conservative 0; Mismatches 514; Indels 6; Gaps 1;
QY 1 AAGATGAAGATGAGACAGTCTTGTGAATCATAACTGTATATCTTGGCCCATTTGGAAC 60
DB 2353 AAGAAGCTGTAAGAGGTGGCTCGGTCTCCATATAGCAATCTTGTCTCATCTTCTGCCAA 2412
QY 61 TGTGGAGGCCCTTAATGATGGTCTCTACTTCCAAATGGTGGCGAAGGATCTGGTTT 120
DB 2413 CCACCTGTACAGTCTCGGTAGACTGTACTATGAATAATGGAGGAAAGCAATATG 2472
QY 121 CCAAGCTCTCTTCAAAACCGTTGGTGTCCCACTCATTTTCTCCCTCTCTCTTATCTTTC 180
DB 2473 GTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGATTTCTCTCGCTTCTTCTCGA 2532
QY 181 CTCGCGCGTGTGGTTCCTTGAAGAACAAAGAACGACTCCATTTTTCCTCATGAACCT 240
DB 2533 ATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCCTTCCCTTCCACCCCTT 2592
QY 241 CTTCTCTTTATCGCCCTATCGTTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 300
DB 2593 GCATCGGTTTAC-----TTGTGCACTGGAGCTAGTGTCTGTCTATGCTTATTTGCT 2646
QY 301 TCTTAGGGTGTAGCTTATATCCCTGTTTCTACTGCGTCTTGTATCATCTCCGCGCAATTA 360
DB 2647 GCAGTTGGGTGTTTACTTACCAAGCTCTACTTTCTCCCTCATCTTGGCCTCACAGTTG 2706
QY 361 GGCCTCACGCTCTCTTTGCATTTTATGGTGAAGCAAAAGTTCACACCTTTTCACTATA 420

DB 2707 GCCTTCACTGCCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTCTTTGTAGATC 2766
QY 421 AAGCGTATCGTTTTCCTCACTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 480
DB 2767 AATTCCTTGTCTTCTTACGGTTTCTCTGCGCTTCTCTGCTGGTCAACACTGATTGAGAA 2826
QY 481 AAGCTTGCAACGAGACACACAAGGAATATGTTGTGGGTTCATCATGACTCTTGGTGA 540
DB 2827 AACACACAAATGTTATCTAGAGTACAAATATGTTGGTTCATCTGTACCACTGGTGT 2886
QY 541 GCTCTTCTATGGGTTTATATTTGACACTTGTGAGCTTCTTACAGAAATCTGGTCA 600
DB 2887 TCCGCTGGGATTTGGTATCTATCTCTGATACAACTGCTCTCAGGAAGTTTTCAGC 2946
QY 601 CGAATCAGTATACGCTGCGCTCGAGTTCAGATGGTCTTATGCTTGTCTGCTGCCACTTGT 660
DB 2947 AAGCATACATCTCTGAGCAGTCTCTGACCTTGGCAATACCAGTCTCTAGTTGCAACTTGT 3006
QY 661 GTCTGCTCTGGGGATGCTAGCCGCTGGCGATTTCAAGGTGATAGCAGGAGAACAAGA 720
DB 3007 GTGGTACTCATAGGACTGTTTGCAGTGGAGAGTGGAGAACTCTGCCAAGTGAGATGAGA 3066
QY 721 GATTTTAACTTGGAGAGTCTTTGTACTATGTTGGTGTGTTGTTTCAAGGCAATATCTGG 780
DB 3067 AACTACAACTGGGGAAAGTGTATATATCTTGAATTTGGGCTCAGCAGCTATTTTCTGG 3126
QY 781 CAAGCATTTTGTGGGAGCTATTGGGTTGATCTTCTGTGATCTGCTGCTGCTGCTGCTGGA 840
DB 3127 CAAGTATACACTGTTGGTGTGGGATTAATCTTCGAGTCTTCTGTGTGTCTTCTTCCAAT 3186
QY 841 ATTATGTCAGTGTCTCTGCTCCGCTGACGGTGTCTTGGCGCTCATTTTGTCTTCCAGGAG 900
DB 3187 TCCATAACTGCTGTGGGACTGCTATAGTTCAGTGTGAGCAGTATAGTTTTCATGAT 3246
QY 901 AAGTTTCAAGCGGGGAAAGGTGTCCTTGGTCTCTCTGCTGCTGGGATCAGTCTCTTAT 960
DB 3247 AAGATGGATGATCAAGATTTTCTCCATCATTTTGTAGTATCTGGGGCTTCTCTTCAATT 3306
QY 961 TTCTATGACAGGTTTAAATCCGAGGAGAAGACTAAGGCTCAGGATACACAAC 1012
DB 3307 GTCTATCAGCACTACTCTCGACGAAAGGTTGAAGACTTGCACAGACAAAAC 3358
RESULT 6
ADA68246
ID ADA68246 standard; DNA; 3387 BP.
XX
AC ADA68246;
XX
DT 20-NOV-2003 (first entry)
XX
DE Arabidopsis thaliana gene, SEQ ID 908.
XX
KW Plant; bacterial infection; fungal infection; viral infection; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

XX Claim 6; SEQ ID NO 908; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC the expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

XX Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;

Query Match 16.0%; Score 167.6; DB 8; Length 3387;
Best Local Similarity 48.6%; Pred. No. 5.7e-42;
Matches 492; Conservative 0; Mismatches 514; Indels 6; Gaps 1;

QY 1 AAGATGAAGATGAAGACAGTCTTGTATATATAAATCTGTATATCTTGGCCATTTGGAAC 60
DB 2353 AAGAACTGAAGAGTGGCTCGGTCTCCATATAGCAATCTTGTGCATCTTCTGCCAA 2412
QY 61 TGTGAGGCCCTTAATGATCGGTCTTACTTCCAAATATGTCGGAAGATCTGGTTT 120
DB 2413 CCACTTGTACAGTTCTGGTAGACTGTACTATGAATAATGGAGGAAAGCACATATGTG 2472
QY 121 CCAAGCTTCCCTCAACCGTGTGTGTCACATCTTCTCCCTCTCTCTTATCTTTC 180
DB 2473 GTAACACTTCTCAACTCATTTGGCTTCCCTGTACTGATCTCTCCGCTTCTTTCGGA 2532
QY 181 CTCGCCCTGCTGTGCTTGAAGACAAAGAACGACTCCATTTTCTCATGAACCT 240
DB 2533 ATCAGGAACCAATCAACAGATACAAATTCAGTCAGTCCCTTCTTACACACCTT 2592
QY 241 CCTCTCTTTATCGCGGCTATCGTTGTGGTTTGTCTGFGGAATTTGACAATTAACCTTAC 300
DB 2593 GCATCGGTTTAC-----TTGTGCACCTGGACTGTAGTGTCTGCTTATGCTTATTTGTCT 2646
QY 301 TCTTACGGGTAGCTTATATCCCTGTTTCTACTCGCTTTGATCATCTCCGCGCAATTA 360
DB 2647 CGAGTTGGTTGCTTTTACAGAGTCTTACTTCTCCCTCATCTTGGCCCTCACAGTTG 2706
QY 361 GGCTTCACTGCTCTTTTGCAATTTTATGGTGAAGCAAAAGTTTCACACCTTTTCACTATA 420
DB 2707 GCCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTTCACCTTTTGATAGTC 2766
QY 421 AACGCTATCGTTTGTCTACCTGCTGGTGCCGTAGTCTTGGCCCTTAACCTGTGATGTGAC 480
DB 2767 AATTCTTTGTTTCTCTTACCGGTTTCTCTGCCCTCTCGTGGTCAACACTGATTCAGAA 2826
QY 481 AAGCTTGCACACGACACACAGGAATATCTGTGGTTTCATCATGACTCTTGTGCA 540
DB 2827 AACCAACAATGATATCTAGAGTCAATATGTGATTTGGTGTTCATCTGTACCAITGGTGCT 2886
QY 541 GCTCTTCTATGGGTTTATATGCGCACTTGTGAGCTTCTTCAAGAAATCTGGTTCAG 600
DB 2887 TCCGCTGGATGGATGGTACTATCTCTGATACACATGCTCTTCAGAAAGTTTTCAG 2946
QY 601 CGAATCAGTATAGCTCGCGCTCGAGTTCAGATGFGTCTTATGCTTTGTGCGCACTGT 660
DB 2947 AAGCATACATCTCAGCAGTCTCGACTTGGCCAAATACCAAGTCTCTAGTTGCAACTGT 3006
QY 661 GTCTGCTCGTGGGATGCTAGCCGCTGGCCATTTCAAGTGTATAGCAGGAGAGCAAGA 720
DB 3007 GTGTACTCATAGGACTGTTTGCAGTGGAGAGTGGAGAACTCTGCCAAGTGAATGAGA 3066
QY 721 GATTTTAAAGCTTGGAGAGTCTTTGTACTATATGTGGTATTTGTTCAGCGCCATAATCTGG 780

DB 3067 AACTACAAATGGGAAAGTGTATATATCTTGACTTTTGGCTCAGCAGCTATTTTCTGG 3126
QY 781 CAAGCATTTTGTGGAGACTATTTGGTTGATCTTCTGTGATCTCTCTGGTCTCTGGA 840
DB 3127 CAAGTATACACTGTTGTGTTGGGATTAATCTTCAGTCTTCTTCTGTGTTCTCCAAT 3186
QY 841 ATTATGTCAGTCTCTGCTTTCGGTGCAGGTGATCTTGCCGTCATTTGCTTCCAGGAG 900
DB 3187 TCCATAACTGCTGGGACTGCTTATAGTTCCAGTTGTAGCAGTATAGTTTTCATGAT 3246
QY 901 AAGTTTTCAGGCGGGAAGGTGTCGTTTGGCTCTCTCCCTCTGGGATCAGTCTCTTAT 960
DB 3247 AAGATGATGCATCAAAAGATTTTCTCCATCATTTTAGCTATCTGGGCTTCTTTCATTT 3306
QY 961 TTTATGACAGGTTAAATCCGAGGAGAACTAAGSCTCAGGATACACAAAC 1012
DB 3307 GTCTATCAGCACTACTCGACGAAAGAGTTGAAGACTTGGCCAGACAAAAC 3358

RESULT 7
AAA97922

ID AAA97922 standard; DNA; 1293 BP.

XX AAA97922;

XX 19-JAN-2001 (first entry)

XX A. thaliana PUP1 DNA #4.

XX PUP1; transgenic plant; nucleobase transporter; apical dominance;
XX flowering behaviour; senescence; pesticide distribution; ds.

XX Arabidopsis thaliana.

XX DE19907209-A1.

XX 24-AUG-2000.

XX 19-FEB-1999; 99DE-01007209.

XX 19-FEB-1999; 99DE-01007209.

XX (FROM/) FROMMER W.

XX Gillissen B, Buerkle L, Andre B, Frommer WB;

XX WPI; 2000-566202/53.

XX Nucleic acid, useful for producing transgenic plants with altered
XX nucleobase transport, encodes a nucleobase transporter protein of
XX Arabidopsis thaliana.

XX Claim 1f; Page 13; 24pp; German.

XX This invention describes a novel nucleic acid encoding a plant nucleobase
XX transporter (I). (I) is produced by complementation of a nucleobase
XX transport (NBT)-defective host cell with a plant gene bank by selection
XX of NBT-positive cells. (I) is used to isolate homologous sequences from
XX bacteria, fungi, plants, animals and humans, for expression of the
XX encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
XX expression of (II) (when in antisense orientation), and to produce
XX transgenic crop plants. The transgenic plants have modified nucleobase
XX transport properties, e.g. altered affinity and substrate specificity
XX that may result in more efficient nucleobase transport in leaves, changes
XX in apical dominance, flowering behaviour and senescence, or improved
XX distribution of pesticides. This sequence encodes the Arabidopsis
XX thaliana PUP1 protein which is described in the method of the invention

XX Sequence 1293 BP; 345 A; 288 C; 247 G; 413 T; 0 U; 0 Other;

Query Match 15.0%; Score 157; DB 3; Length 1293;
Best Local Similarity 50.5%; Pred. No. 8e-39;
Matches 379; Conservative 0; Mismatches 371; Indels 0; Gaps 0;

Db 667 ATCCCGCTCGTGGAGTCCACGCGGAGGTACGGGCGCGCAGCGGGCCCGCGCGAGG 726
Qy 606 -----CAGGTATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTTGCTGCCACT 657
Db 727 GTGCCGCTCCCTACGCGACGGTGATGAGATGACAGGCGGTGATGGGCGCGGCGGACG 786
Qy 658 TGTGTCGCTCGTGGGATGCTAGCCGCTGGCGATTTCAAGGTGATAGCAGGAGCA 717
Db 787 GCGGTGCTCGTCTCGGATGCGCATCAAGGCGGACTTCCAGGCGGTGGCGCGGAAGCG 846
Qy 718 AGAGATTTTAAGCTTGGAGAGTCTTTGTACTATGTGGTGATTGTCTTCAAGGCGCATATC 777
Db 847 GCGGCTTTCGGGCTCGGCGCGGCGCACTACTACCTCGTCTCGCTGGGACGCGGTGCG 906
Qy 778 TGGCAAGCATTTTGTGGGAGCTATTGGGTGATCTTCTGTGATCGTCTCTGCTCTCT 837
Db 907 TGGCAGCTGCTCAACCTGGGCATCATGGGCTCATCACCTGCGGCTGCTGCTGCTGCC 966
Qy 838 GGAATTATGCTAGTCTCTGCTCCGCTGAGGTGATCTTGGCGGTCATTTGCTTCCAG 897
Db 967 GGCATCATGATGCGCGTCTCTCGCGCTCTCGCAGGTCTCTCGCGCTCATCTTCTCCAC 1026
Qy 898 GAGAAGTTTTCAGGCGGGAAGGTGCTGCTTTGGCTCTCTCCCTCTGGGATCAGTCTCT 957
Db 1027 GAGAAGTTTCAGCGGACGAGGCAATCGGCTGCTGCTCTGCTCTGGGATTCGCTCC 1086
Qy 958 TATTCTATGACAGTTTAAATCCGAGGAGAG 990
Db 1087 TACTCTACGCGGAGGCGGAGAGAGAG 1119

RESULT 9

AAA97925
ID AAA97925 standard; DNA; 1071 BP.
AC AAA97925;
XX
DT 19-JAN-2001 (first entry)
DE A. thaliana PUP1 DNA #7.
XX
XX PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.
XX Arabidopsis thaliana.
OS
XX DE19907209-A1.
XX
XX PD 24-AUG-2000.
XX
XX PF 19-FEB-1999; 99DE-01007209.
XX
XX PR 19-FEB-1999; 99DE-01007209.
XX
XX PA (FROM/) FROMMER W.
XX
XX PI Gillissen B, Buerkle L, Andre B, Frommer WB;
XX
XX DR WPI; 2000-566202/53.
XX
XX PT Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transport, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.
XX
XX PS Claim 2; Page 15; 24pp; German.
XX
XX CC This invention describes a novel nucleic acid encoding a plant nucleobase
CC transporter (I). (I) is produced by complementation of a nucleobase
CC transport (NBT)-defective host cell with a plant gene bank by selection
CC of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
CC expression of (II) (when in antisense orientation), and to produce

CC transgenic crop plants. The transgenic plants have modified nucleobase
CC transport properties, e.g. altered affinity and substrate specificity
CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX
SQ Sequence 1071 BP; 307 A; 236 C; 208 G; 320 T; 0 U; 0 Other;
Query Match 14.7%; Score 154.2; DB 3; Length 1071;
Best Local Similarity 49.6%; Pred. No. 5.6e-38;
Matches 458; Conservative 0; Mismatches 453; Indels 12; Gaps 2;
Qy 49 GCCATTGGAACTGCTGGAGCCCTCTAATGATGGCTCTCTACTTCCAAAATGGTGGCGAA 108
Db 46 GTCAATGGCCAAATCAGTTGCTACAAATTCGGGCAGACTATACTATGAAAATGGAGGAAC 105
Qy 109 AGGATCTGGTTTCCAGCTTCCCAACCGTTGCTTCCACTCATTTCTTCCCTCTT 168
Db 106 AGCAATGGCTTAGCAACGGTAGTTTACGCTTGTAGGCTTCTCTATTTCTACTTCCATATCAT 165
Qy 169 CTCTTATCTTTCTCCCGCTCGTCTGCTTGAAGAACAAAGAAAGAGTCCATTTTTC 228
Db 166 CTCTTGTCTAGTCAAAACACAT-----ACAACTCAGAGAGATGCAAAATTAAC 215
Qy 229 CTGATGAAACCTCTCTCTTTATCGCGCTATGCTTGGTGTGCTGCGGATTTGAC 288
Db 216 CTCACCTTAGGAACCGTGCAATTA--GTTTACATATGCTTGGACTTCTTGTAGGAGAGCT 273
Qy 289 AATTACCTTACTCTTACGGGTTAGCTTATATCCCTGTTTCTACTGCTCTTTGATCATC 348
Db 274 TGTACCTATATTCATTGGACTGCTTACCTACTGTTTCTACCTTTTCCCTGATCTGT 333
Qy 349 TCCGCGCAATTAGGCTTCACTGCTCTCTTTGCAATTTTTTATGTTGAGCAAAAGTTCACA 408
Db 334 GCATCAGATTAGCTTTCACCGCTTCTCTCTTATTTACTCAACTCACAAGAACTTACT 393
Qy 409 CTTTCACTATAAACGCTATCGTTTTCCTCACTGGTGGTGGCTGCTTCCCTTAAAC 468
Db 394 CCTATCATTTTGAATTTCTTTTCCCTCACTATATCTTCCACCTCTTGGCATTTAAT 453
Qy 469 TCTGATAGTCAGAGCTTGCAGAGAGACACAGGAATATGTTGTTGGTTCATCATG 528
Db 454 AACGAGGAATCAGATTTCCAAAAGTTTCAAAAGGAGATGTCAAAGGTTTCGTATGC 513
Qy 529 ACTCTTGGTGCAGCTCTTCTATGGGTTTATATGCTTATGCTTGTGAGCTTCTTCAAG 588
Db 514 ACCGTTGGTGCATCTGCTGGGTTTGGTCTACTCTTATCCCTACAAAGCTAGCCTTTCGT 573
Qy 589 AATCTGCTCAGGAATCAGTATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTT 648
Db 574 AAAGTTTTAAAGAGCAAACTTTCTCAGAAAGTTATAAATATGATAATCTACATGAGTCTA 633
Qy 649 GCTGCCACTTGTGCTCGCTCGGGGATGCTAGCGCTCGAGTTCAGATGGTCTTATGCTTT 708
Db 634 GTGCCAGTTGTGTAGCTGGTGGGCTTTTCTAGTAGGAGTGGAAGAACTTTGAGC 693
Qy 709 GGAGAAGCAAGAGATTTTAAAGCTTGGAGAGTCTTTGTACTATGTTGGTGTATGTTCAAG 768
Db 694 AGTGAATGGAAAACTACAAACTTTGGGAAGGTATCTTATGTCATGAACCTAGTGTGACA 753
Qy 769 GCCATATCTGGCAGCAATTTTGTGGGAGCTATTTGGGTTGATCTTCTGTGCATCTCT 828
Db 754 GCTGTTACCTGGCAGGTATTTCTCCATCGGTTGACAGGAGTGAATCTTTCGAGCTTCTCTCC 813
Qy 829 CTGCTCTCTGGAATTAAGTTCAGTCTCTGCTTCCGGTACGGTGAATCTTGGCGCTCAT 888
Db 814 CTATTTCTCAATGCAATGACGGCTTTGGGACTCTCCCGTGGTTCCTATCTTGGCTGTATC 873
Qy 889 TGCTTCCAGGAGAGTTTTCAGGCGGGAAGGTTGCTGCTTGGCTCTCTCTCCCTCTGGGA 948
Db 874 ATTTTCCATGACAAAATGAACGGCTTAAAGGTGATTCTTATGATTTCTAGCTATTTGGGGT 933

QY 949 TCAGTCTCTTATTTCTATGACA 971
Db 934 TTGCTATCCTATGCTCTACCAACA 956

RESULT 10
AAC37652
ID AAC37652 standard; DNA; 1403 BP.
XX
AC AAC37652;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 18172.
XX
KW Hybridisation assay; Genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.

CC	thaliana PUP1 protein which is described in the method of the invention	
XX		
SQ	Sequence 1081 BP; 322 A; 231 C; 205 G; 323 T; 0 U; 0 Other;	
	Query Match 14.2%; Score 148.8; DB 3; Length 1081;	
	Best Local Similarity 50.6%; Pred. No. 3e-36;	
	Matches 360; Conservative 0; Mismatches 352; Indels 0; Gaps 0;	
QY	260 TCGTGTGTTGTTGTCGTGGGATTTGACATTAACCTCTACTCTTACGGGTAGCTTATA	319
DB	254 TAGTGTCTGGACTTCTTGTAGAGCAGATTCCTATCTGTATCCATTGCACTTCTTACT	313
QY	320 TCCCTGTTCTACTGCGCTTTTGATCATCTCGGCAATTAGGTTCACTGCTCTCTTTG	379
DB	314 TACCGTTTACCTATTCCCTGATCTGTGATCTCAGTTAGCCTTCAATGCTTCTTCT	373
QY	380 CATTTTTATGGTGAAGCAAAAGTTTCAACCTTTCACTTATAAAGCTATCGTTTTGTCA	439
DB	374 CTTATTTTCTTAACCTACAAAACCTTACCCTATCATTTTAAATTTCTTTTCTCTTAA	433
QY	440 CTGTTGGTGGCTAGTCTTCCCTTAACCTGATAGTGAAGCTTGCACAAAGACAC	499
DB	434 CTATATCTTCCACCCCTACTTGCATTTCAATAATGAGGACAGACTCCCAAAAGTTACA	493
QY	500 ACAAGGAATATGTTGTTGGTTTCATCATGACTCTTGTGCAGCTCTTCTCTATGGTTTA	559
DB	494 AAGGAGATGTTCAAGGTTTTCATGACCGTTGCTGCTGCTGCTGTTATGTTCTAG	553
QY	560 TATTGCCACTTGTGAGCTTTCTTACAAGAAATCTGFTCAGCAATCACGTATACGCTCG	619
DB	554 TCTTATCCCTACAACAGCTAGCCTTTCTAAAAGCTCTAAAGCAAAATTTCTCAGAAG	613
QY	620 CGCTCGAGTTCAGATGTTCTATGCTTTGTCGCACTTGTGCTGCTGCTGCTGGGATGC	679
DB	614 TTATGGATATGATATCTACGTGAGTCTAGTGGCCAGTTGTGTTAGCGTGGTGGGCTTT	673
QY	680 TAGCGCTGGCGATTTCAAGGTGATAGCAGGAGAAAGATTTTAAAGCTTCGAGAGT	739
DB	674 TTGCTAGCAGTGAAGTGGAACCTTTGACGATGAATGGATACTACAAACATGGGAAG	733
QY	740 CTTTGTACTATGTGTGATGTTGTTCACGGCCATAATCTGGCAAGCAATTTTTTGTGGAG	799
DB	734 TATCTACATTTATGAACCTAGTGTGGACAGCTTTACCTGGCAGTATTTCCCATCGTG	793
QY	800 CTATTGGTTCATCTTCTGTGCATCGTCTCTGGTCTCTGGAAATATGTCAGTCTCTGC	859
DB	794 GCACAGGACTGATCTTCGAGCTCTCCTCTCTATTCTCAAAATGCAATAGCGTTTTGGGAC	853
QY	860 TTCGGTGACGGTGAATCTTGGCCGTCATTTGCTTCCAGGAGAAAGTTTCAGGCGGGAAAG	919
DB	854 TCCAGTGGTTCCTATCTTGGCTGTATCATTTTCCATGACAAATGATGGGTTAAAG	913
QY	920 GTGTGCTTTGGCTCTCTCCCTCTGGGATCAGTCTCTTTATTTCTATGGACA	971
DB	914 TGAATTTCTATGATCTAGCTATTTGGGGTTTCACTTCTCTATGTTCTCAACA	965
RESULT 13		
AAC47974		
ID	AAC47974 standard; DNA; 2660 BP.	
XX		
AC	AAC47974;	
XX		
DT	18-OCT-2000 (first entry)	
XX		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 55799.	
XX		
KW	Hybridisation assay; genetic mapping; gene expression control;	
KW	protein identification; signal transduction pathway; metabolic pathway;	
XX	promoter; termination sequence; ss.	
OS	Arabidopsis thaliana.	
XX		

PN	EPI033405-A2.	
XX		
FD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-00301439.	
XX		
PR	25-FEB-1999; 99US-0121825P.	
PR	05-MAR-1999; 99US-0123180P.	
PR	09-MAR-1999; 99US-0123548P.	
PR	23-MAR-1999; 99US-0125788P.	
PR	25-MAR-1999; 99US-0126264P.	
PR	29-MAR-1999; 99US-0126785P.	
PR	01-APR-1999; 99US-0127462P.	
PR	06-APR-1999; 99US-0128234P.	
PR	08-APR-1999; 99US-0128714P.	
PR	16-APR-1999; 99US-0129845P.	
PR	21-APR-1999; 99US-0130077P.	
PR	19-APR-1999; 99US-0130449P.	
PR	23-APR-1999; 99US-0130510P.	
PR	23-APR-1999; 99US-0130891P.	
PR	28-APR-1999; 99US-0131449P.	
PR	30-APR-1999; 99US-0132048P.	
PR	30-APR-1999; 99US-0132407P.	
PR	04-MAY-1999; 99US-0132484P.	
PR	05-MAY-1999; 99US-0132485P.	
PR	06-MAY-1999; 99US-0132486P.	
PR	06-MAY-1999; 99US-0132487P.	
PR	07-MAY-1999; 99US-0132863P.	
PR	11-MAY-1999; 99US-0134256P.	
PR	14-MAY-1999; 99US-0134218P.	
PR	14-MAY-1999; 99US-0134219P.	
PR	14-MAY-1999; 99US-0134221P.	
PR	14-MAY-1999; 99US-0134370P.	
PR	19-MAY-1999; 99US-0134768P.	
PR	19-MAY-1999; 99US-0134941P.	
PR	20-MAY-1999; 99US-0135124P.	
PR	21-MAY-1999; 99US-0135353P.	
PR	24-MAY-1999; 99US-0135629P.	
PR	25-MAY-1999; 99US-0136021P.	
PR	27-MAY-1999; 99US-0136392P.	
PR	28-MAY-1999; 99US-0136782P.	
PR	01-JUN-1999; 99US-0137222P.	
PR	03-JUN-1999; 99US-0137528P.	
PR	04-JUN-1999; 99US-0137502P.	
PR	07-JUN-1999; 99US-0137724P.	
PR	08-JUN-1999; 99US-0138094P.	
PR	10-JUN-1999; 99US-0138540P.	
PR	14-JUN-1999; 99US-0138847P.	
PR	16-JUN-1999; 99US-0139119P.	
PR	16-JUN-1999; 99US-0139452P.	
PR	16-JUN-1999; 99US-0139453P.	
PR	17-JUN-1999; 99US-0139492P.	
PR	18-JUN-1999; 99US-0139454P.	
PR	18-JUN-1999; 99US-0139455P.	
PR	18-JUN-1999; 99US-0139456P.	
PR	18-JUN-1999; 99US-0139457P.	
PR	18-JUN-1999; 99US-0139458P.	
PR	18-JUN-1999; 99US-0139459P.	
PR	18-JUN-1999; 99US-0139460P.	
PR	18-JUN-1999; 99US-0139461P.	
PR	18-JUN-1999; 99US-0139462P.	
PR	18-JUN-1999; 99US-0139463P.	
PR	18-JUN-1999; 99US-0139750P.	
PR	18-JUN-1999; 99US-0139763P.	
PR	21-JUN-1999; 99US-0139817P.	
PR	22-JUN-1999; 99US-0139899P.	
PR	23-JUN-1999; 99US-0140353P.	
PR	23-JUN-1999; 99US-0140354P.	
PR	24-JUN-1999; 99US-0140695P.	
PR	28-JUN-1999; 99US-0140823P.	
PR	29-JUN-1999; 99US-0140991P.	
PR	30-JUN-1999; 99US-0141287P.	
PR	01-JUL-1999; 99US-0141842P.	

Dbb 890 AAGCTACACTTGTGCTTGTGGGATTTGATCTTCGAGTCATCCCTCTGTGTTCTCCAATT 949
Qy 842 TTATGCTCAGTCTCTGCTTCCGGTGAACGGTGATCTTGGCCGTCATTTGCTTCCAGGAGA 901
Db 950 CCATACAGCTGTGGGATTCCTATAGTCCAGTTCGGGCGAGTAGATTTTCCAGATA 1009
Qy 902 AGTTTCAGCGGGGAAAGGTGCTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTATT 961
Db 1010 GAATGGACGATCAAAAATCTTCTCCATTTATTTAGCTATCTCGCGCTTCCTTTCATTCG 1069
Qy 962 TCTATGACAGGTTAAATCCGAGGAGAACTAAGGCTCAGGATACA 1008
Db 1070 TCTATGACACTACCTCGACGAAAGATGTAATACTAGCCACACA 1116

RESULT 15

ABN85767

ID ABN85767 standard; cDNA; 83698 BP.

AC ABN85767;

XX 21-OCT-2002 (first entry)

XX Arabidopsis yellow stripe1-like 4 encoding cDNA SEQ ID NO 9.

DE Maize; transgenic; plant; yellow stripe1-like; ysl; Arabidopsis;
KW iron uptake; bioremediation; yellow stripe 1; ysl; gene; ss.

XX Arabidopsis sp.

XX WO200240688-A2.

XX 23-MAY-2002.

XX 16-NOV-2001; 2001WO-US043101.

XX 16-NOV-2000; 2000US-0249222P.

XX (UYVA) UNIV YALE.

XX Walker EL, Dellaporta S;

XX WPI; 2002-490144/52.

DR P-PSDB; AB883920.

XX New yellow stripe1 and yellow stripe1-like genes, useful for altering the
PT distribution of iron within the plant body so that edible parts of crop
PT plants have more iron, or for producing plants useful in enhancing iron
PT uptake from soil.

XX Claim 1; Page 114-159; 187pp; English.

XX The invention relates to an isolated nucleic acid molecule (I), maize
CC yellow stripe 1 (ysl) or yellow stripe1-like (ysl) from Arabidopsis
CC (ABN85763-ABN85771). (I) is useful for generating transgenic plants which
CC can be used for enhancing iron uptake from soil and for bioremediation of
CC metal or heavy metal contaminated soil. (I) may also be used to alter the
CC distribution of iron within the plant body so that edible parts of crop
CC plants have more iron. Transgenic plants may also be used in conventional
CC plant breeding schemes to produce progeny which also contain the gene of
CC interest. The present sequence is that of the Arabidopsis ysl encoding
CC cDNA of the invention

SQ Sequence 83698 BP; 26295 A; 15150 C; 15104 G; 27149 T; 0 U; 0 Other;

Query Match 12.7%; Score 133; DB 6; Length 83698;

Best Local Similarity 48.0%; Pred. No. 3.8e-30;

Matches 455; Conservative 0; Mismatches 475; Indels 17; Gaps 2;

Qy 20 TTCTGTATATCAATACTGATATATCTTGGCCATTGGAACTGTGGAGGCCCTCTAATGA 79

Db 67235 TTCTGTGTTTCAATAGCATCTTCTTCTCATCTCTGCTCAAGCCATTCTCTTCTCTTG 67294

Qy 80 TCGCTCTCTACTTCCAAAATGGTGGCAAGGATCTGTTTCCAAAGCTTCTCTTCAAACCG 139
Db 67295 GTCCGTTTTATTACAAATGAAGGTGAAACAGTAATGGATCTCTACTCTTGTTCAAACGTG 67354
Qy 140 TTGG-TTGTCCACTCAATTTCTTCCCTCTCTCTTATCTTTCTCCGCGCTCGTGTGCG 198
Db 67355 GTGGCTTTCCCAATCTTTATCTCCCTCTTCTCTCTCTTCTCTGCTTCCAAATCTTCTCTT 67414
Qy 199 CTTGAAGAAACAAGAAACGACTCCATTTTCTCATGAAACCTCTCTCTTTATTCGCGGT 258
Db 67415 CTT-----CTTCTTCTTCTTCTTCTTCAAGACTCTGGTTTGGATTAT 67458
Qy 259 ATCGTTGTGTTGCTCTCGGGATTTCACAATTAACCTACTCTTACGGGTAGCTTAT 318
Db 67459 CTTTCTCTTGGTTTGTCTATTGGTTAGATAAATTTTATATACTCTGTGGACTTTTGTAT 67518
Qy 319 ATCCCTGTTTCTACTGCGCTCTTGTATCATCTCCGGCGCAATTAGGCTTCACTCTCTCTT 378
Db 67519 CTTCTCTGTTCAACTTATTCGAATCTATGTGCTTCCACAGTTAGCTTTCAATGGTGTCTTC 67578
Qy 379 GCATTTTTTATGGTGAAGCAAAAGTTACACCTTTTCACTATATAACGCTATCGTTTGTCTC 438
Db 67579 TATTAATACATCAATCTCAGAAAATCACTTGTTTGATTTTCTTCTCAGTGTGTCTCTC 67638
Qy 439 ACTGTTGTCGCGTAGTCTTTCGCCCTTAACTCTGTAGTGAACAAGCTTGCAAAACGAGACA 498
Db 67639 TCTATCTCTGCTGTGTTGTTTCTCTTGCAGATGATTCAAATAGCCCATCAGGAGATTCT 67698
Qy 499 CACAAGGAATATGTTGTTGGGTTTCATCATGACTCTTGGTGCAGCTCTTCTCTATGGGTTT 558
Db 67699 AAGTGGAGTTACTTGAATGGGTGTTTCTGTGCAGTTTTTGTCTTCTTATCTATCTCTT 67758
Qy 559 ATATTGCCACTTGTGCGAGCTTTCTTACAAGAAATCTGTTCAAGCAATCACGTTATACGCTC 618
Db 67759 CAGCTCTCTCTATGACAGTTTTTCTTTCAGAAAGTTCTCAAGAGTGAGACTCTCTCTATG 67818
Qy 619 GCGTCGAGTTCAGATGCTCTTATGCTTGTGCTGCCACTTGTGTCTGCTCTGCTGGGATG 678
Db 67819 GTTCTCGAGATGCAAAATCTATAGCTGCTTGTGGCTTCTTGTGTAGCGGTATTCGATTG 67878
Qy 679 CTAGCCGCTGGCGATTTCAGGTGATAGCAGGAGAAAGCAAGAGATTTTAAGCTTGGAGAG 738
Db 67879 TTCCGAAGCGGGGAATGGATGTTGTAGTGTGGAGATGGAAGAGTTTTCAGGAAGTCAA 67938
Qy 739 TCTTTGTACTATGTGGTGATTGTTTACGGCCATAATCTGCAAGCAATTTTTCGTGGGA 798
Db 67939 GTCAATTAATGTTTGAATTTGCTGCGGCGAGCGGTTTCTGTCATATTCGGTTGTGTAGA 67998
Qy 799 GCTATTTGGGTTGATCTTCTGTGCATCGTCTCTGGTCTCTGGAATTTATGGTCTCTGCTG 858
Db 67999 GCGGTGTCGTTATATTTCTGTTGCTTTCGCTGTTTTCAAAACCTTATTTAGTACGCTCTCA 68058
Qy 859 CTTCCGTTGACGGTGATCTTGGCGCTCATTTGCTTCCAGGAGAAAGTTTTCAGCGGGGAAA 918
Db 68059 CTCATTTGTACGCTCTCCTCGCGCCATTTCGGGTGTTCATGATGCAAGCTGACTGAGGTTAAG 68118
Qy 919 GGTGTCGCTTGGGCTCTCTCCCTCTGGGGATCAGTCTCTTATTTCTTA 965
Db 68119 ATGTTGGCGATGCCCATCGCCTTACAGGATTTACGTTTATATCTTA 68165

Search completed: November 1, 2004, 07:19:42

Job time : 559.844 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query			DB	ID	Description
		Match	Length				
c	1	1049	100.0	1049	6	AX033545	AX033545 Sequence
	2	997.4	95.1	1075	8	AF078532	AF078532 Arabidops
	3	701.8	66.9	114144	8	U78721	U78721 Arabidops1s
	4	528	50.3	1225	6	AX033544	AX033544 Sequence
	5	528	50.3	1227	8	AF078531	AF078531 Arabidops
	6	526.4	50.2	1250	8	AK117664	AK117664 Arabidops
	7	519	49.5	1102	8	BT005504	BT005504 Arabidops
	8	478	45.6	1067	6	AX033553	AX033553 Sequence
	9	408.2	38.9	121720	8	AC021044	AC021044 Arabidops
	10	251.6	24.0	1505	8	AK066610	AK066610 oryza sat
	11	242	23.1	1458	8	AK061667	AK061667 Oryza sat
	12	226	21.5	1363	8	AK064135	AK064135 Oryza sat
	13	167.6	16.0	1165	8	AY036558	AY036558 Arabidops
	14	167.6	16.0	1428	8	AY074546	AY074546 Arabidops
	15	167.6	16.0	3387	6	AX507517	AX507517 Sequence
	16	167.6	16.0	3387	6	AX652015	AX652015 Sequence
	17	167.6	16.0	82697	8	ATT9A21	AL021713 Arabidops
	18	167.6	16.0	194143	8	ATHCH1V48	AL161548 Arabidops
	19	163.6	15.6	73179	8	AP006383	AP006383 Lotus cor
	20	163.6	15.6	73179	8	AP006383	AP006383 Lotus cor

Db	181	CTCCGCGTGGTGGCTTGAAGACAAAGAGCACTCCATTTTCTCATGAACCT	240
Qy	241	CCTCTCTTTATCGCGCTATCGTTGGTTGGTTGCTCGTGGGATTGCAATTAACCTCTAC	300
Db	241	CCTCTCTTTATCGCGCTATCGTTGGTTGGTTGCTCGTGGGATTGCAATTAACCTCTAC	300
Qy	301	TCCTACGGGTTAGCTTATATCCGTTTCTACTGGCTTTTGATCATCTCCGCGCAATTA	360
Db	301	TCCTACGGGTTAGCTTATATCCGTTTCTACTGGCTTTTGATCATCTCCGCGCAATTA	360
Qy	361	GGCTTCACTGCTCTCTTTCATATTTTATGGTGAAGCAAAAGTTTACACCTTTTCACTATA	420
Db	361	GGCTTCACTGCTCTCTTTCATATTTTATGGTGAAGCAAAAGTTTACACCTTTTCACTATA	420
Qy	421	AACGCTATGTTTGGTCTCACTGGTGGTCCGAGTCTTTCAGAAATCTGGTCAG	480
Db	421	AACGCTATGTTTGGTCTCACTGGTGGTCCGAGTCTTTCAGAAATCTGGTCAG	480
Qy	481	AAGCTTGCAACGAGACACACAGGAATATGTTGTTGGGTTTCACTGACTCTTGGTGCA	540
Db	481	AAGCTTGCAACGAGACACACAGGAATATGTTGTTGGGTTTCACTGACTCTTGGTGCA	540
Qy	541	GCTCTTCTATGGTTTATATGTCACCTGCTCGAGCTTCTTACAGAAATCTGGTCAG	600
Db	541	GCTCTTCTATGGTTTATATGTCACCTGCTCGAGCTTCTTACAGAAATCTGGTCAG	600
Qy	601	CGAATCACTATACGCTCGCGCTCGAGTTCAGATGCTTATGCTTTGTCGCCACTGT	660
Db	601	CGAATCACTATACGCTCGCGCTCGAGTTCAGATGCTTATGCTTTGTCGCCACTGT	660
Qy	661	GCTGCTCGTGGGATGCTAGCGCTGGCAATTTCAAGGTGATAGCAGGAGAGCAAGA	720
Db	661	GCTGCTCGTGGGATGCTAGCGCTGGCAATTTCAAGGTGATAGCAGGAGAGCAAGA	720
Qy	721	GATTTAAAGCTTGAGAGCTCTTGTACTATGTTGATGTTGTTGTTGTTGTTGTTGTTG	780
Db	721	GATTTAAAGCTTGAGAGCTCTTGTACTATGTTGATGTTGTTGTTGTTGTTGTTGTTG	780
Qy	781	CAAGCATTTTGTGGGAGCTATTGGGTGATCTTCTGTGATCGTCTCTGGTCTCTGGA	840
Db	781	CAAGCATTTTGTGGGAGCTATTGGGTGATCTTCTGTGATCGTCTCTGGTCTCTGGA	840
Qy	841	ATTATGTCAGTCTGCTCTCCGCTGACGTGATCTTGGCCGTCATTTGCTTCCAGGAG	900
Db	841	ATTATGTCAGTCTGCTCTCCGCTGACGTGATCTTGGCCGTCATTTGCTTCCAGGAG	900
Qy	901	AAGTTTCAAGCGGGGAAGGTGTGCTTTGCTCTCTCCCTCTGGGATCAGTCTCTTAT	960
Db	901	AAGTTTCAAGCGGGGAAGGTGTGCTTTGCTCTCTCCCTCTGGGATCAGTCTCTTAT	960
Qy	961	TTCTATGACAGGTTAAATCCGAGAGAACTAAGGCTCAGGATACAACTGTCTCAG	1020
Db	961	TTCTATGACAGGTTAAATCCGAGAGAACTAAGGCTCAGGATACAACTGTCTCAG	1020
Qy	1021	CTTCCAGTTACTGATTATGACTTAAA	1049
Db	1021	CTTCCAGTTACTGATTATGACTTAAA	1049

RESULT 2
AF078532
LOCUS AF078532
DEFINITION Arabidopsis thaliana putative purine permease (PUP2) mRNA, partial cds.
ACCESSION AF078532
VERSION AF078532.1
KEYWORDS GI:14388590
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE	1 (bases 1 to 1075)	
AUTHORS	Gillissen,B., Burkle,L., Andre,B., Kuhn,C., Rentsch,D., Brandl,B. and Frommer,W.B.	
TITLE	A new family of high-affinity transporters for adenine, cytosine, and purine derivatives in Arabidopsis	
JOURNAL	Plant Cell 12 (2), 291-300 (2000)	
MEDLINE	20129770	
PUBMED	10662864	
REFERENCE	2 (bases 1 to 1075)	
AUTHORS	Gillissen,B., Andre,B., Rentsch,D., Buerkle,L., Kuehn,C., Brandl,B. and Frommer,W.B.	
TITLE	Direct Submission	
JOURNAL	Submitted (15-JUL-1998) Institut fuer Botanik, University of Tuebingen, Auf der Morgenstelle 1, Tuebingen, Baden-Wuerttemberg 72076, Germany	
FEATURES	Location/Qualifiers	
source	1..1075	
	/organism="Arabidopsis thaliana"	
	/mol_type="mRNA"	
	/cultivar="C 24"	
	/db_xref="taxon:3702"	
	/chromosome="II"	
gene	1..>1075	
	/gene="PUP2"	
CDS	1..>1075	
	/gene="PUP2"	
	/note="AlPup2"	
	/codon_start=1	
	/product="putative purine permease"	
	/protein_id="AAK61813.1"	
	/db_xref="GI:14388591"	
	/translation="MKMKTVLVIINCFIAGNCGPLMMRLYFQNGGERIMWPSFLQ TVGCLIFPFLLSFURRRLEQETTPFLMKPFIAAIVVGLVDFNDLYSYG LAYIPVSTASLIISAQGLTALFAFFWVKQFTPTTINAILVTGGAVLALNSDSK LANETHKEYVVGFIIMTLGALYGFILPLVELSYKSGQRITTYTALFQMVLCAAT CGLVGMLAGDFKVKHALFIFKNRVIAGEARDFKLGESLYYVVIPTAIWQAFFVG AIGLIFCASSLVSIMVSALLPVTIVILVICFQEKFAQGKVALALSWSVSFYFQQ VKSEETKAQDTQLSQLPVTDYA"	
ORIGIN		
	Query Match 95.1%; Score 997.4; DB 8; Length 1075;	
	Best Local Similarity 96.8%; Pred. No. 7.6e-289;	
	Matches 1041; Conservative 0; Mismatches 1; Indels 33; Gaps 1;	
Qy	4	ATGAAGATGAAGACAGTCTCTTAATCATFAAATCTATATCTTGGCCATGGAACTGT 63
Db	1	ATGAAGATGAAGACAGTCTCTTAATCATFAAATCTATATCTTGGCCATGGAACTGT 60
Qy	64	GGAGGCCCTCTAATGATGCTCTCTACTTCCAAAATGGTGGCGAAAGGATCTGGTTCCA 123
Db	61	GGAGGCCCTCTAATGATGCTCTCTACTTCCAAAATGGTGGCGAAGGATCTGGTTCCA 120
Qy	124	AGCTTCCTTCAAAACCGTTGGTTGTCCACTCAITTTCTTCCCTCTTCTTATCTTCTC 183
Db	121	AGCTTCCTTCAAAACCGTTGGTTGTCCACTCAITTTCTTCCCTCTTCTTATCTTCTC 180
Qy	184	CGCGTCTGCTGGTTCCTTGAAGAACAGAAACGACTCCATTTTCTCATGAACCTCT 243
Db	181	CGCGTCTGCTGGTTCCTTGAAGAACAGAAACGACTCCATTTTCTCATGAACCTCT 240
Qy	244	CTCTTTATCGCGCTATCGTTGTTGGTTGCTCGTGGGATTTGACAAATTAACCTCTACTCT 303
Db	241	CTCTTTATCGCGCTATCGTTGTTGGTTGCTCGTGGGATTTGACAAATTAACCTCTACTCT 300
Qy	304	TACGGGTAGCTTATATCCCTGTTTCTACTCGCTCTTTGATCATCTCCGCGCAATTAGGC 363
Db	301	TACGGGTAGCTTATATCCCTGTTTCTACTCGCTCTTTGATCATCTCCGCGCAATTAGGC 360
Qy	364	TTCACTGCTCTCTTGCATTTTATATGGTGAAGCAAAAGTTGCACACCTTTCACTATAAC 423
Db	361	TTCACTGCTCTCTTGCATTTTATATGGTGAAGCAAAAGTTGCACACCTTTCACTATAAC 420
Qy	424	GCTATCGTTTTGTCTCACTGGTGGTGGCTAGTCTTGGCCCTTAACCTCTGATGATGACAAG 483

mRNA
CDS
repeat_region
gene
mRNA
CDS
repeat_region
repeat_region
gene
mRNA
CDS
gene
mRNA
CDS

join(<7252..7434,7522..7713,7864..8069,8275..9045)
/gene="At2g33700"
join(7252..7434,7522..7713,7864..8069,8275..9045)
/gene="At2g33700"
/codon_start=1
/product="putative protein phosphatase 2C"
/protein_id="AAC69126.1"
/db_xref="GI:1707015"
translation="MSMDFSLLTVLEGDFNKNONTSSATEIDTLENLDDTROISKGKP
PHLTSATRIQLAANDVDVNLVMSKLDKSEFLPVYRSGSCAEQAKQFWEDEHI
CIDLVNHLGAIAOCSSILGAFYGVDFDGGTDAAHFKNIIIRFIVEDSFPPLCVKKA
IKSAFLADFEADDSSLDISSGTALTARIFGRPLIIANAGDCRAVLGRGRIELS
KDHPCNTAEKVIKELGGVYDGYLNGQLSVARAIGDWMKPKGSACPLSPPELQ
ETLSEDDFEILMGCDGVYDMSQCAVTTIARKELMIHNDPERCSRSLREALKRNIC
DNLTVIVCFSPDPFRIETRMQSRVRSISAEGNLNLLKGLVDGP"
complement(11621..11649)
/rpt_family="AT rich"
complement(12323..13704)
/gene="At2g33710"
note="synonym: TlB8.3"
complement(join(<12323..12798,13524..>13704))
/gene="At2g3710"
complement(join(12323..12798,13524..13704))
/gene="At2g33710"
/codon_start=1
/product="putative AP2 domain transcription factor"
/protein_id="AAC69127.1"
/db_xref="GI:1707016"
translation="MHSGKRPLSPESMAGNRREKKELCCSTLSESDVDFVSELTCQ
PIPSIDDOSSSLTLOKSNRSRVRGVQRPWKWAAIRIDPNKAAKRVMLGTFDTA
EEAALVDKAAFERGHAKLNFFEHIRVAPTQLYPSPATSHDRIIIVTPSPPPPIAP
DILIDQYGHFOSRSSDSANLNMNLSSSSSSLNHQGLRNLNEDGENVKNISIHKKRK
"
complement(14124..14156)
/rpt_family="AT rich"
complement(14389..14425)
/rpt_family="TA)n"
16828..18041
/gene="At2g33720"
note="synonym: TlB8.29; predicted by genscan and
genefinder"
join(<16828..17139,17373..>18041)
/gene="At2g33720"
join(16828..17139,17373..18041)
/gene="At2g33720"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC69145.1"
/db_xref="GI:2253580"
translation="MKRPPPSASKTQYLFDHDESSNSKSLVSTTLLSLSCENPNKR
KNSDEVNLNISCIPRDYKLTQVERKIAMRNLSYOEKADEWYGVSTELTLFKDPWII
KKVFFASVLDMAPNSVNTLCLDTPESPENAEESLSDLCFYDKTWPDPNVAYNK
PTSEAINLAWMTMSKRAREEKYVSTELTLLTVADPWLKAMTKSSIGNLYRL
VLKASFDVIDIHLRPLDDOMMKEDSGLAVEVYDHTDTSVHNIALKKWAKSSSFVLV
SGWRKCFVDRGLQVDVIGMYDRSBSKLHFCVLSRSETMDSAPLPSP"
complement(18731..20932)
/gene="At2g33730"
note="synonym: TlB8.4"
complement(<18731..>20932)
/gene="At2g33730"
complement(18731..20932)
/codon_start=1
/product="putative U5 small nuclear ribonucleoprotein, an
RNA helicase"
/protein_id="AAC69128.1"
/db_xref="GI:1707017"
translation="MKRTINIEVSGTSVNTLDDLSIAKKPVFLTKKQREELAKRQDQ
ISEQRVREOIGRPEPETEDVSGDNTNRDNRDRDRDRDRDRDRDRDRDRDRDR
DR
DAIKEQYLGKKPKKRVIRPSEKFRFDFMENTEDTSRDNMVLYQNPEHAQLLFGRF
RAGMDREOKQAAKHEKMERDEIRKKDGI VEKPEEAAAKRVEEAAADTVDSFDMRVYD
RHWSDKRLBEETWDRIFREDFNI SYKSGRI PRPMRSWEESKLITSELLKAVERAGYK

KPSPIOMAAIPLGLQORDVIGIAETSGKTAFAFVLPMLAVLSRLPMPMSENETEGPYA
VWNPARELAQIEEBETKFAHYLIGFRVTSIVGGQSTEBQGLKITQCEIVIAITPGRLL
IDCLERRYAVLNQCNVYVLDDEADMDMGPEQVAGVLDMPSSNLKPENEEBELDK
KIYRTTYMVSATMPGPVERLARKYLNRNVVVTITGAKTTDLISQHVIMKSEKFPFR
KYNLVDLGBKTAIVFVNTKKNCDSTAKNLDKAGYRVTLTGKGKSQBQRISLSEGPRA
KYNLVAITDVVGRGIDIPDAHVINDMPKHIEMVTHRIGRTGRAGKSGVATSLFLTL
HDTVEPYDLKQMLVQNSAVPPELARHEASRFKPGTVDPDRPHSDTVYIN"
complement(20588..20718)
/rpt_family="(CAGAGA)n"
complement(21321..22493)
/gene="At2g33735"
note="synonym: TlB8.17; supported by full length cDNA:
Ceres: 269036"
complement(join(<21321..21655,21744..21814,22051..22126,
22221..22298,22384..>22493))
/gene="At2g33735"
complement(join(21584..21655,21744..21814,22051..22126,
22221..22298,22384..22446))
/gene="At2g33735"
/codon_start=1
/product="Expressed protein"
/protein_id="AAM15518.1"
/db_xref="GI:20198310"
translation="MMSDWDGSDHGFDFDFEDYKDHYKVLNCDASDDEIRSSPFR
LALKWHPDRFKEDSATSRPQENEAQVLSDFIAQOEYDKRMRIYEHNTLLNEY
LKLILTCNLGLGVNHYLM"
22524..22544
repeat_region
Query Match 66.9%; Score 701.8; DB 8; Length 114144;
Best Local Similarity 99.7%; Pred. No. 1.9e-199;
Matches 703; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 AAGATGAAGATGAAGACAGTCTTGTAATCATAAATGTATATCTTTGGCCATTTGGAAC 60
Db 26320 AAGATGAAGATGAAGACAGTCTTGTAATCATAAATGTATATCTTTGGCCATTTGGAAC 26321
Qy 61 TGTGGAGGCCCTCTAATGATGCGTCTCTACTTCCAAATCGTGGCGAAGATCTGTTT 120
Db 26260 TGTGGAGGCCCTCTAATGATGCGTCTCTACTTCCAAATCGTGGCGAAGATCTGTTT 26201
Qy 121 CCAAGCTTCCTTCAAACCGTTGGTTGTCCACTCATTTTCTTCCCTCTTCTTATCTTTC 180
Db 26200 CCAAGCTTCCTTCAAACCGTTGGTTGTCCACTCATTTTCTTCCCTCTTCTTATCTTC 26141
Qy 181 CTCGCCGTCGTGTCGTTGCCTTGAAGAACAGAACGACATCCATTTTCTCATGAACCT 240
Db 26140 CTCGCCGTCGTGTCGTTGCCTTGAAGAACAGAACGACATCCATTTTCTCATGAACCT 26081
Qy 241 CCTCTCTTTATCCCGCTATCGTTGTTGTTGCTGCTGGGATTTGACAAATTACCTCTAC 300
Db 26080 CCTCTCTTTATCCCGCTATCGTTGTTGTTGCTGCTGGGATTTGACAAATTACCTCTAC 26021
Qy 301 TCTTACGGGTAGCTTATATCCCTGTTTCTACTCGCTCTTTTGATCATCTCGCGCAATTA 360
Db 26020 TCTTACGGGTAGCTTATATCCCTGTTTCTACTCGCTCTTTTGATCATCTCGCGCAATTA 25961
Qy 361 GGCTTCACTGCTCTCTTTGCAATTTTTTATGGTGAAGCAAAAGTTACACCTTTTACCTATA 420
Db 25960 GGCTTCACTGCTCTCTTTGCAATTTTTTATGGTGAAGCAAAAGTTACACCTTTTACCTATA 25901
Qy 421 AACGCTATCGTTTGTGCTCACTGGTGGCGGTAGTCTTCCCTTAACCTCTGATAGTGAC 480
Db 25900 AACGCTATCGTTTGTGCTCACTGGTGGCGGTAGTCTTCCCTTAACCTCTGATAGTGAC 25841
Qy 481 AAGCTTCAACAGAGACACAAAGGAATATGTTGTTGGGTTTCATCATGCTCTTGGTGCA 540
Db 25840 AAGCTTCAACAGAGACACAAAGGAATATGTTGTTGGGTTTCATCATGCTCTTGGTGCA 25781
Qy 541 GCTCTTCTATAGGTTTATATGCGCACTTGTGAGCTTTCTTACAAGAAATCTGGTCAG 600
Db 25780 GCTCTTCTATAGGTTTATATGCGCACTTGTGAGCTTTCTTACAAGAAATCTGGTCAG 25721
Qy 601 CGAATCAGTATACGCTCGCGCTCGAGTTCAGATGCTCTTATGCTTTTGTGCTCCACTTGT

```
|||||
Db 25720 CGAATCAGTATACGTCGCGCTCGAGTTCAGATGCTCTATGCTTGTGCGCACTTGT 25661
QY 661 GTCTGCTCGTGGGATGCTAGCGCTGGCGATTTCAGAGTGATA 705
Db 25660 GTCTGCTCGTGGGATGCTAGCGCTGGCGATTTCAGAGTAAAA 25616

RESULT 4
AX033544 1225 bp DNA linear PAT 21-SEP-2000
LOCUS
DEFINITION Sequence 1 from Patent WO049152.
ACCESSION AX033544
VERSION AX033544.1 GI:10280288
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Andre,B., Buerkle,L., Frommer,W.B. and Gillissen,B.
TITLE Nucleic acids that code for a nucleobase transporter
JOURNAL Patent: WO 0049152-A 1 24-AUG-2000;
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;
GILLISSEN BERND (DE)
FEATURES
source
1. .1225
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 50.3%; Score 528; DB 6; Length 1225;
Best Local Similarity 70.8%; Pred. No. 2.2e-147;
Matches 762; Conservative 0; Mismatches 285; Indels 33; Gaps 3;

QY 1 AAGATGAAGATGAAGACAGTCTTGTAAATCATAACTGTATATCTTTGGCCATGGAAAC 60
Db 16 AAGAAGAGATGAAGATGGTTTGATATATCATAACTGTATATCTCTACTATAGAACA 75
QY 61 TGTGGAGGCTCTAATGATGCTCTACTATCCAAATGGTGGGAAAGGATCTGGTTT 120
Db 76 TGTGGAGGCTCTTGTAACTCTCTACTTACCAGTGGCGGAAAGGAAATCTGGTTC 135
QY 121 CCAAGCTTCCTCAACCGTGGTGTGCCACTATTTCTTCCCTCTCTCTATCTTTC 180
Db 136 ATGAGCTTCTATCAACCGTGGTGTTCCAATCATCTCATCCCTCTCTTGGTCTCTTC 195
QY 181 CTCGCGCGTCTGCTGTG-----CCTTGAAGAACAAAGAAACGACTCCA 222
Db 196 CTCAGCGTCTCGCGGCAACCGCAACCTTAACAACGCGGAAACAAAGCAAAAG 255
QY 223 TTTTCTCTATGAACACTCTCTCTTTATACGCGCTATCGTTGTTGGTTGCTGGGA 282
Db 256 CTCTCTCTATGAACACTCTCTCTTTATCGCTCCATGTCATAGGTTGCTCACAGGA 315
QY 283 TTTGACATTTACTCTACTCTTACGGGTAGCTATATCCCTGTTTCTACTGCGCTTTC 342
Db 316 CTTGACAATCTATATATCTTACGGATAGCATATCTGCGAGTTTCAACTTCATCGCTC 375
QY 343 ATCATCTCCGCGCAATTAGCTTCACTGCTCTCTTTGCAATTTTATGGTGAAGCAAAAG 402
Db 376 ATATCGGAATCACTAGCTTTCAACGCTCTCTTCTGCTTTCTTGTAGTCAAGCAAAAG 435
QY 403 TTCACACTTCTACTATAAAGCTATCGTTTTCCTCACTGGTGGTCCGCTAGTCTTGGC 462
Db 436 TTCACCTCGTCTCCATAAAGCGCGTCTTTTGTGACGGTGGTATCGGATCCTTGGC 495
QY 463 CTTAACTCTGATGATGACAGCTTGCAGAGACACACACAGCAATATGTTGGGTTC 522
Db 496 TTACACAGTGTGAGACAAACCGGCTAAGGAGCAAGAAAGAGTATGTGGTGGTTC 555
```

```
QY 523 ATCATGACTTTGGTGACGCTCTTCTCTATGGGTTTATATGCCACTTGTGCGAGCTTTCT 582
Db 556 TTGATGACTGTGTTGCAGCTCTTCTCTATGCTTTTATATACCGCTCGTTGAGCTAACT 615
QY 583 TACAAGAAATCTGGTCAGCAATCAGTATACGCTCGCGCTCGAGTTCAGATGGTCTTAA 642
Db 616 TACAAGAAAGCTCGTCAAGAAATCACTTTCCACATTGCTTGAGATTGAGATGGTCAATG 675
QY 643 TGCTTTGCTGCCACTTGTGCTGCGCTCGTGGGATGCTAGCGCTCGCGATTTTCAAGGTG 702
Db 676 TGCTTTGCTGCTACTTTTCTGTCATTTGSCATGTTTCATGTTGGAGATTTTAAGGTG 735
QY 703 ATAGCAGGAGAAAGCAAGAGATTTTAAAGCTTGAGAGTC---TTTGTACTATGTGGTATT 759
Db 736 ATAGCAAGAGAAAGCAAGAGATTTCAAGATTTGAGGATCAGTGTGTTTACTATGCAATTGATA 795
QY 760 GTGTTACCGGCCATATCTGGCAAGCATTTTGTGGGAGCTATTGGGTTGATCTTCTGT 819
Db 796 GTGATCACAGGAATTAATATGGCAAGGTTTCTTCTTAGGAGCCATAGGGATTTGTGTTTGT 855
QY 820 GCATCGTCTCTGCTCTCTGGAATTATGTCAGTGTCTCTCCGCTGACGGTGATCTTGT 879
Db 856 GCATCATCACTAGTCTCTGCTGTTCTGATAAGTGTCTGCTTCCGGTGACTGAAGTTTTC 915
QY 880 GCGTCATTTGCTTCAGGAGAAAGTTTCAGGCGGAAAGGTTGCTTGGCTCTCTCTCC 939
Db 916 GCGTCTGTTGTTTCCGGGAGAAAGTTTCAGGCAAGAGTGTCTCTCTCTACTTCTTCT 975
QY 940 CTCTGGGATCAGTCTCTTATTTCTATGGACAGGTTAAATCC-----GAGGAG 987
Db 976 CTTTGGGATTTGCTCTCTTACTTCTACGCGGAGTTTAAATCCGGCAAGAAAGTTGTGAT 1035
QY 988 AAGACTAAGGCTCAGGATACACAACTGCTCAGTTCCTCAGTTCCTCAGTATGTAGCTTAA 1047
Db 1036 AAACCTCAACCCCGGAGACAGAACTGCTATCTTCCAGTAGTATGATGTTGCTTAA 1095

RESULT 5
AF078531 1227 bp mRNA linear PLN 20-APR-2000
LOCUS
DEFINITION Arabidopsis thaliana purine permease (PUP1) mRNA, complete cds.
ACCESSION AF078531
VERSION AF078531.1 GI:7620006
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Gillissen,B., Buerkle,L., Andre,B., Kuhn,C., Rentsch,D., Brandl,B.
and Frommer,W.B.
TITLE A new family of high-affinity transporters for adenine, cytosine,
and purine derivatives in Arabidopsis
JOURNAL Plant Cell 12 (2), 291-300 (2000)
MEDLINE 20129770
PUBMED 10662864
REFERENCE
AUTHORS Gillissen,B., Andre,B., Rentsch,D., Buerkle,L., Kuehn,C., Brandl,B.
and Frommer,W.B.
TITLE Direct Submission
JOURNAL Submitted (15-JUL-1998) Institut fuer Botanik, University of
Tuebingen, Auf der Morgenstelle 1, Tuebingen, Baden-Wuerttemberg
72076, Germany
FEATURES
source
1. .1227
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
1. .1227
/gene="PUP1"
27. .1097
CDS
```

		/gene="PUP1" /note="AtPup1" /codon_start=1 /product="purine permease" /protein_id="AAF64547.1" /db_xref="GI:7620007" /translation="MKNGLIILINCIILITIGTCGGPLLRLYFTNGKRIWPMFLSTA GFPIILILVLSFSSRRGRNPNNAENKRTKPLMETPLFIASIVIGLITGLDNYL YSYGLAYLPVSTSSLIIGTQAFNALFAFLVKKQFTPFSSINAVLLNVGIGLALHS DGDKPKSKKEYVGVPLMTVVAALLYAFILPLVELTYKKARQELTFPLVLEIQWMC LAATPFVIGHFIVGDFKVIAREAREPKIGGSVFYIALIVITIGLIWQGFAGIVF CASSIAGSLVSLVLPVTEVFAVCFREKFORAEKGVSLLSLNGVFSVYFGEKRSKK VVDKPPPETELPILPVS DYVA"	
ORIGIN			
Query Match		50.3%; Score 528; DB 8; Length 1227;	
Best Local Similarity		70.6%; Pred. No. 2.2e-147;	
Matches 762; Conservative 0; Mismatches 285; Indels 33; Gaps 3;			
Qy	1	AAGATGAAGATGAAGACAGTCTCTGTAATCATAACTGTATATTTCTTGGCCATTGGAAC	60
Db	18	AAGAAGAAGTGAAGATGGTTTGATATATCATAACTGTATATCTCACTATAGGAACA	77
Qy	61	TGTGAGGCCCTTAAATGATGGCTCTTACTTTCAAATATGGTGGCGAAGATCTGGTTT	120
Db	78	TGTGAGGTCTCTTTTAACTCGTCTCTACTTCACCAATGGCGGAAAACGAATCTGGTTC	137
Qy	121	CCAAGCTTCTTCAACCGTGGTTGTCCACTCATTTTCTCCCTCTCTCTTATCTTTC	180
Db	138	ATGAGCTTCTTATCAACCGTGGTTTCCAAATCCTCATCCCTCTCTTGGTCTCCCTTC	197
Qy	181	CTCCGCGTCTGGTGTG-----CCTTGAAGAACAAACGACTCCA	222
Db	198	CTCAGCGTCTCGCGGGCAACCGCAACCTTAACAGCGGGAACCAACGCGGAACAAG	257
Qy	223	TTTTTCTCTGAACACTCTCTCTTTATCGCGGTATCGTGTGTGTTCTCGTGGGA	282
Db	258	CTCTTCTCATGGAACCTCTCTGTTTATCGCTCTCAATGTTCATAGGTTGCTCACAGGA	317
Qy	283	TTTGACAATTACTCTACTCTTACGGGTAGCTTATATCCCTGTTTCTACTGCGCTTTG	342
Db	318	CTTGACAATCTTATATCTTACGGATAGCATATCGCCAGTTTCAACTTCATCGCTC	377
Qy	343	ATCATCTCGCGCAATTAGGCTTCACTGCTCTCTTTTGCAATTTTATGGTGAAGCAAAAG	402
Db	378	ATAATCGGAATCAACTAGCTTCAACGCTCTCTTCGCTTCTTGTGTAGTCAAGCAAAAG	437
Qy	403	TTACACCTTTTCACTATAAAGCTATCGTTTGTCTCACTGGTGGTGGCGTATGCTTGGCC	462
Db	438	TTCACTCGGTTCTCCATAAAGCGGCTCGTTTGTGTGACGGTTGGTATCGGGATCCTTGGC	497
Qy	463	CTTAACCTGATAGTGACAAGCTTGCAACGACAGACACACAGGAATATGTTGGGTTTC	522
Db	498	TTACACAGTATGAGACAAACCGGCTAAGGAGAGCAAGAAAGATATGTTGGGTTTC	557
Qy	523	ATCATGACTCTGTGTGAGCTCTTCTCTATCGGTTTATATTGCCACTTGTGCGACTTCT	582
Db	558	TTGATGACTGTGGTTGCGAGCTCTTCTATGCTTTTATATTACCGCTCGTTGAGCTAACT	617
Qy	583	TACAAGAAATCTGGTACCGAATCAGGTATACGCTCGCGCTCGAGTTCAGATCGTCTTA	642
Db	618	TACAAGAAGCTCGTCAAGAAATCACTTCCCACTTGTGCTTGAGATTTCAGATGGTCATG	677
Qy	643	TGCTTTGCTGCACCTTGTGTGCTGCTCGTGGGATGCTAGCCGCTGGCGATTTCAAGGTG	702
Db	678	TGCCTTGTGCTACTTTTTCTGTGTCATGGCAATGTTTCATCGTTGAGATTTTAAAGGTG	737
Qy	703	ATAGCAGGAGAAGCAAGAGATTTTAAAGCTTGGAGATC---TTTGACTATGTGGTGATT	759
Db	738	ATAGCAAGAGAAGCAAGAGAGTTCAGATTGGAGATCAGTGTTTTACTATGCAATGATA	797
Qy	760	GTGTTTCAGGCCATAATCTGGCAAGCAATTTTCTTGTGGAGAGCTATTTGGGTGATCTTCTGT	819
Db			
		1038 AACTCTCAACCGCGGAGACAGAACTGCCTATTCTTCCAGTTAGTATATTGTTGCTTAA	
		1097	
RESULT 6			
AKI17664			
LOCUS			
DEFINITION		Arabidopsis thaliana Atlg28230 mRNA for putative purine permease,	
complete cds, clone: RAF117-32-017.			
AKI17664			
ACCESSION			
VERSION		AKI17664.1 GI:26450404	
KEYWORDS		FLI CDNA; CAP trapper.	
SOURCE		Arabidopsis thaliana (thale cress)	
ORGANISM		Arabidopsis thaliana	
		Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.	
REFERENCE			
AUTHORS		Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,	
		Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,	
		Kawai, J., Hayashizaki, Y. and Shinozaki, K.	
		Arabidopsis thaliana full-length cDNA	
		Published Only in Database (2002)	
TITLE		2 (bases 1 to 1250)	
JOURNAL		Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,	
AUTHORS		Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,	
		Kawai, J., Hayashizaki, Y. and Shinozaki, K.	
		Direct Submission	
		Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences	
		Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa	
		230-0045, Japan (E-mail: meeki@gsc.riken.go.jp).	
		URL: http://pfweb.gsc.riken.go.jp, Tel: 81-45-503-9625,	
		Fax: 81-45-503-9586	
		An Arabidopsis full-length cDNA library was constructed essentially	
		as reported previously (Seki et al. (1998) Plant J. 15:707-720;	
		Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI	
		and XhoI was ligated to modified Lambda FIC-1-E vector (Carninci et	
		al. (2001) Genomics 77:79-90) digested with BamHI and SalI.	
		This clone is in a modified pBluescript vector.	
		Please visit our web site (http://pfweb.gsc.riken.go.jp/) for	
		further details.	
FEATURES		Location/Qualifiers	
source		1. 1250	
		/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/db_xref="taxon:3702"	
		/chromosome="1"	
		/clone="RAF117-32-017"	
		/ecotype="Columbia"	
		/notes="common name: thale cress"	
gene		1. 1250	
CDS		53..1123	
		/gene="Atlg28230/F3H9_10"	
		/codon_start=1	
		/product="putative purine permease"	
		/protein_id="BAC42317.1"	
		/db_xref="GI:26450405"	

gene
CDS
3' UTR
ORIGIN

/clone="U50269"
/ecotype="Columbia"
/note="This clone is in pUNI 51."
1..1102
/gene="At1g28230"
1..11071
/gene="At1g28230"
/codon_start=1
/evidence=experimental
/product="putative purine permease"
/protein_id="RA063924.1"
/db_xref="GI:28973199"
/translation="MNGLIILINCIILFTGCGPLRLRYFTNGGKRIWFMFLSTAFPIILPLVSLFRSRENPNNAENKRLKFLMETPLFIASIVIGILGLDNYLVSYLGLVPTSSLIIGTQAFNALPFLVVKOKFTPFSSINAVVLTVGIGILALHS DGDPAKSKKSYVGVVTVVALLVAFILPLVELTYKKARQBITPLVLEIOWMC LAAFPFCVIGMFVIGDFKVIAREAREFKIGGSVPYALIVITIGIQGFFGAGIGIV CASSIAGSVLISVLVPTVEFVAVVCVREKFOAEKGVSLLSLWGFVSFYFGEFKSGK VVDPKPPETELPLIPVSDYA"
1072..1102
/gene="At1g28230"

Query Match 49.5%; Score 519; DB 8; Length 1102;

Best Local Similarity 70.3%; Pred. No. 1.1e-144;

Matches 753; Conservative 0; Mismatches 285; Indels 33; Gaps 3;

QY 10 ATGAAGACAGTCTTGTATCATAACTGTATATTCTTGGCCATTGGAACTGTGGAGGC 69
DB 1 ATGAAGAATGGTTGATTAATCAATAACTGTATTATCTCACTATAGGAACATGTGGAGGT 60
QY 70 CTTCTAATGATCGCTCTACTTCCAAATGGTGGCGAAAGGATCTGGTTTCCAAAGCTTC 129
DB 61 CTTTGTGTTAACTGGTCTCTACTTCACCAATGGCGGAAACGAATCTGGTTCTATGAGCTTC 120
QY 130 CTTCAACCGTGGTGTCCACTCATTTTCTCCCTCTCTCTTATCTTCTTCTCTCCGCGGT 189
DB 121 CTATCAACCGTGGTGTTCCAATCATCTCATCCCTCTCTTGGTCTCTCTCTCCAGCGGT 180
QY 190 CGTCGTTG-----CCTTGAAGAACAGAAACGACTCCATTTTCTCTC 231
DB 191 CGCGCAGACACCGCAACCTTAACACCGCGGAAACAAACGCGGAAACAAAGCTCTTCTC 240
QY 232 ATGAACCTCTCTCTTTATCGCGCTATCGTTGTGTTGTTGCTCGTGGGATTTGACAAT 291
DB 241 ATGGAACCTCTCTGTTTATCGCTCCATTTGTCATAGGTTGCTCACAGGACTTGACAAC 300
QY 292 TACCTCTACTCTTAGCGGTAGCTTATATCCCTGTTTCTACTGCGTCTTTGATCATCTCC 351
DB 301 TACTTATATCTTACGGATTAGCATATCTGCCAGTTTCAACTTTCATCGCTCATATCGGA 360
QY 352 GCGCAATTAGGCTTCACTGCTCTCTTTGCAATTTTATGTTGTAAGCAAAAGTTCCACCT 411
DB 361 ACTCACTAGCTTTCAACGCTCTCTTCGCTTCTGTTAGTCAAGCAAAAGTTTCACTCG 420
QY 412 TTCACTATAAACGGCTATCGTTTGTCTACTGCTGGTGGCGTAGTCCCTTGGCCCTTAACTCT 471
DB 421 TTCTCCATAACGGCGTGTGTTTGTGACGGTTGGTATCGGATCCCTTGGCTTACACAGT 480
QY 472 GATAGTGACAGCTTGCAGACACAGACACAAAGGAATAGTTGTTGGTTCATCATGACT 531
DB 481 GATGAGACAAACCGGCTTAAGGAGAGCAAGAAAGATATGTGTTGGGTTCTTGATGACT 540
QY 532 CTTGGTGAGCTCTTCTCTATGGGTTTATATTTGCCACTTGTGCAGCTTTCTTACAAGAA 591
DB 541 GTGTTGAGCTCTTCTCTATGCTTTTATATACCGCTCGTTGAGCTTAACATTACAAGAA 600
QY 592 TCTGGTCAGCAATCACGTATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTTGCT 651
DB 601 GCTCGTCAAGAAATCACTTTCCCACTTGTGCTTGAGATTACAGATGGTCAATGCTTGTCT 660
QY 652 GCCACTTGTCTCGCTCGTGGGATGCTACCGCTGCGGATTTCAAGGTGATAGCAGGA 711

DB 661 GCTACTTTTCTGTGTCATTTGGCAATGTTTCATCGTTGGAGATTTTAAAGTGATAGCAAGA 720
QY 712 GAAGCAAGAGATTTTAAGCTTTGGAGAGTC---TTTGTACTATGTGGTGAATGTGTTACG 768
DB 721 GAAGCAAGAGAGTTTCAAGATTGGAGGATCAGTGTTTACTATGCAATTGATAGTATACACA 780
QY 769 GCCATAATCTCGCAAGCATTTTGTGGGAGCTATTTGGGTTGATCTTCTGTGCATCGTCT 828
DB 781 GGAATAATATGCAAGGTTTCTTCTTAGGACCATAGGATTTGTGTTGTGCATCATCA 840
QY 829 CTGCTCTCTGGAATATATGTCAGTGCTCTGCTCCGCTGACGGTGACGGTGATCTTGGCGGTCAAT 888
DB 841 CTAGCTTCTGCTGTTTCTGATAAGTGTCTGCTCCGCTGACTGAAGTTTTCGCCGCTCGTT 900
QY 889 TGCTTCCAGGAGAAGTTTTCAGCGCGGGAAGGTCGCTGTTGGCTCTCTCCCTCTGGGGA 948
DB 901 TGTTCGCGGAGAAGTTTTCAGGCGAGAGAAGGTGTCTCTACTTCTTCTTCTTTGGGGA 960
QY 949 TCAGTCTCTTATTTCTTATGGACAGGTTAAATCC-----GAGGAGAAGACTAAG 996
DB 961 TTTGCTCTTACTTCTACGGCGAGTTTAAATCCGGCAAGAAAGTTGTTGATAAACCTCAA 1020
QY 997 GCTCAGGATACACAACCTGTCTCAGCTTCCAGTTACTGATTATGATGAGCTTAA 1047
DB 1021 CCGCGGAGACAGAACTGCCTATTCTTCTGTTAGTGATTATGTTGCTTAA 1071

RESULT 8
AX033553 LOCUS AX033553 1067 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 10 from Patent WO0049152.
ACCESSION AX033553
VERSION AX033553.1 GI:10280295
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
Andre,B., Buerkle,L., Frommer,W.B. and Gillissen,B.
Nucleic acids that code for a nucleobase transporter
Patent: WO 0049152-A 10 24-AUG-2000;
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;
GILLISSEN BERND (DE)
FEATURES
source
1..1067
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Query Match 45.6%; Score 478; DB 6; Length 1067;
Best Local Similarity 68.1%; Pred. No. 2.5e-132;
Matches 680; Conservative 0; Mismatches 315; Indels 3; Gaps 1;

QY 2 AGATGAAGATGAACAGAGTCTTCTGTAATCATAACTGTATATCTTCGCCATTTCGAAACT 61
DB 1 AGACAGAATGGTGAAGGCTCTTGTGATCAATACTGCAATAATCTAGCCATAGAAACT 60
QY 62 GTGAGAGCCCTCTAATGATCGTCTCTACTTCCAAAAATGGTGGCGAAAGGATCTGGTTTC 121
DB 61 GTGAGAGCTCTTGTATTATGCGTCTCTACTTCAACATATGGCGGTAAAGGATTTGGTTCT 120
QY 122 CAAGCTTCTCTCAACCGTTGGTTGTCACATCTTCTCCCTCTTCTCTATCTTCTCC 181
DB 121 CTAGCTTCTTGAACACTGCAGGCTTTCCCTGTATCTTCAATTCCTCTCTCTCTTCTTACA 180
QY 182 TCCGCGCTCGTCTGTTGCTTTGAAGAACAGAAACGACTCCATTTTCTCTCATGAAACCTC 241
DB 181 TTACCGCGCGAGAGCAACATGTGGTGATAGTACAGTTCTTCTTCTTATCAACGCGC 240
QY 242 CTCTCTTTATCGCGGCTATCGTTGTTGTTGCTCGTGGGATTTGACAAATACCTTACT 301

Db 241 GTCTCTTATCGCGCGTGTATTGTAGGCATTTCTCAGGGTTTGTAACTACTCTGTATG 300
Qy 302 CTTACGGGTTAGCTTATATCCCTGTTTCTACTCGTCTTTTGATCATCTCCGCGCAATTAG 361
Db 301 CATATGGTATAGCTTATCTTCAGTGTCTTACAGCTGCTTATCATTTGCTTACAGTTAG 360
Qy 362 GCTTCACTGCTCTCTTTGCACTTTTATGGTGAAGCAAAAGTTTACACACCTTTTCACTATAA 421
Db 361 CTTTATATAGCTATCTTCTCATTTCTCATGTTTAAACATAAAGTTTCACTCTTTTACCATCA 420
Qy 422 ACCTATCTTTTGTCTGCTGCTGGTGGCGGTAGTCTTGCCTTAACTCTGTATGTAGACA 481
Db 421 ATGCTGTGTGTTGTGACTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 480
Qy 482 AGCTTGCAACAGGAGACACACAGGAATATGTTGTTGGTTCATCATGACTCTTGTGGTGCAG 541
Db 481 AGCAGTTCTATGAGACTCACACAGTACATAACTGTTTCTTGATTTACTGTAGCAGCAG 540
Qy 542 CTCTTCTATGGGTTTATATTCACCTTGTCCAGCTTTTTCACAGAAATCTGGTTCAGC 601
Db 541 CTGTTATGATGCTTTTTCATCTTCCATTTAGTGAACCTTGTACACAGAAAGCTTAACCAA 600
Qy 602 GAATCAGTATACGCTCGCGCTCGAGTTCAGATGCTTATATGCTTGTGCTGTCACCTTGTG 661
Db 601 CCATGAGCTATACCTTGTGCTCGAGTTCAGTGTGATTTGTGCTCTCTTGTCTTCTATTG 660
Qy 662 TCTGCTCTGTTGGGATGCTAGCGCTGGGATTTCA ---AGTGATAGCAGGAGAGCAA 718
Db 661 TCAGCTCATCGTATGTTTCATCGCTGGTGAATTTCAAGCAGGCTTACCAAAAGAGCAA 720
Qy 719 GAGATTTAAGCTTGGAGAGTCTTTGTACTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 778
Db 721 GAGAGTTCAAGCTTGGAGAGCATTGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 780
Qy 779 GGCAAGCATTTTGTGGAGCATTTGGTGTGATCTTCTGTGATCGTCTGCTGCTGCTGCTG 838
Db 781 GGCAAGCTTCTTCTTGGAGCATTGGATTAATCTTCCACATCGTCTGCTGCTGCTGCTG 840
Qy 839 GAATTAATGTCAGTCTCTGCTTCCGTCGAGCGTATCTTGGCGTCATTGTTGTTCCAGG 898
Db 841 GTATTAATGATCAGTCTTTTCCCAATACAGAGTTTGTAGCTGTATATCTACCATG 900
Qy 899 AGAAGTTTCAGGCGGGAAGTGTGCTGTTGCTCTCTCCCTCTCGGGATCAGTCTCTT 958
Db 901 AAAAGTTTCAAGCTGAGAGGACTTCTCTGCTCTCTCCCTTGGGCTTGTCTCTTCT 960
Qy 959 ATTTCTATGACAGGTTAAATCCGAGGAGAACTAAG 996
Db 961 ACTTTTATGTCAGATAAAGTCTGGCGAGGATAAAGG 998

RESULT 9

AC021044 121720 bp DNA linear PLN 30-OCT-2002
LOCUS Arabidopsis thaliana chromosome I BAC F3H9 genomic sequence,
DEFINITION complete sequence.

ACCESSION AC021044

VERSION AC021044.5 GI:8347959

KEYWORDS HTG.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 121720)

Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Mitanda,M., Brooks,S.,
Buehler,E., Chao,O., Chin,C., Chiu,J., Choi,E., Gonzalez,A.,
Hwang,B., Johnson-Hopson,C., Khan,S., Kim,C., Koo,T., Lee,J.M.,
Lenz,C., Liu,A., Liu,S., Mukharshy,N., Pham,P., Sakano,H.,
Shinn,P., Tortumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J.,
and Davis,R.W.

REFERENCE

AUTHORS

FEATURES
source

1..121720
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"

JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 121720)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E.,
Chin,C., Chiu,J., Choi,E., Dunn,P., Gonzalez,A., Hwang,B., Kim,C.,
Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,
Mukharshy,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P.,
Thaveri,A., Tortumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J.,
Theologis,A. and Davis,R.W.
Direct Submission
Submitted (12-JAN-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

TITLE
JOURNALREFERENCE
AUTHORS

3 (bases 1 to 121720)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E.,
Chin,C., Chiu,J., Choi,E., Dunn,P., Gonzalez,A., Hwang,B., Kim,C.,
Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,
Mukharshy,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P.,
Thaveri,A., Tortumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J.,
Theologis,A. and Davis,R.W.
Direct Submission
Submitted (29-MAR-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

TITLE
JOURNALREFERENCE
AUTHORS

4 (bases 1 to 121720)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E.,
Chin,C., Chiu,J., Choi,E., Dunn,P., Gonzalez,A., Hwang,B., Kim,C.,
Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,
Mukharshy,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P.,
Thaveri,A., Tortumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J.,
Theologis,A. and Davis,R.W.
Direct Submission
Submitted (08-JUN-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

TITLE
JOURNALREFERENCE
AUTHORS

5 (bases 1 to 121720)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A.
and Davis,R.W.
Direct Submission

TITLE
JOURNAL

Submitted (11-AUG-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

REFERENCE
AUTHORS

6 (bases 1 to 121720)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A.
and Davis,R.W.
Direct Submission

TITLE
JOURNAL

Submitted (15-AUG-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

COMMENT

On Jun 8, 2000 this sequence version replaced gi:7340332.
Bases 1-7,373 of clone F3H9 overlap with bases 96,792-104,163 of
IGF BAC clone F3M18. gb|AC010155
e-mail for correspondence: arabesequence.stanford.edu Genes with
similarity to proteins in the databases are named 'putative',
'-like' or 'similar to'. Genes that have EST similarity but no
significant protein similarity are described as 'unknown proteins'.
Genes that are annotated based only on gene prediction software
are described as 'hypothetical proteins'. The gene prediction
programs used to predict genes include: Grail (Informatics Group,
Oak Ridge National Laboratory,
http://combio.ornl.gov/section/index.html), GENSCAN (Chris Burge,
http://genome.stanford.edu/~chris/GENSCAN.html), and NetPlantGene
(S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/NetPlantGene.html).

Location/Qualifiers

1..121720

/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"

```
/chromosome="1"
/misc_feature
1..7373
/note="Overlap with bases 96,792-104,163 of IGF BAC clone
F3M18, gb|AC010155, see GenBank record for clone F3M18 for
annotation in this region."
8952..10816
/gene="F3H9.1"
/note="synonym: DRM1"
join(9852..9957,10049..10274,10780..10816)
/gene="F3H9.1"
/note="Location of cDNA clone gb|AF053746 and genomic
clone gb|AF053747."
/codon_start=1
/product="dormancy-associated protein"
/protein_id="AAF98422.1"
/db_xref="GI:9795604"
/translation="MVLLEKLDWDDVAGPQDRGLGRKIKTIPINIRDIGESSK
VMSRLTMPAAVSPGTPTTPTTPRKDNVRSVFNPGSNLATRAIGSNI PDKPTH
VNSPSVDWLYSGDSRSQHR"
complement(12229..14900)
/gene="F3H9.2"
complement(join(12229..12924,13048..13089,13397..13665,
14024..14447,14511..14900))
/gene="F3H9.2"
/note="Hypothetical protein"
/codon_start=1
/protein_id="AAF98443.1"
/db_xref="GI:9795625"
/translation="MDSRDMLACEDEIYSIKNETSLVDLPFGIKPIGLKWFKLIK
RNSVDINKYKARLVAKGYVORYGTDFVEVPVACIEKI ILVDLEASNDWEIOLDVK
TAFILGELKFTVYTRDEYVEKEGKLNQILOPEKCSKEPSVYKTIISGEFLL
VAVYDDLFTVGTWTTIIDEFGEMTSNFMNSNKGKTYILGIEVFQKBEIRLNQSR
YEMILKEAWSCSLNHSRESVVQLAKSNEKEIDATGFRKKVGCRLRYLLHLLSE
VTGLSKRWIRIRIINKSAISLTKNPVPFHGECHKHRRYHFTIRVENEQIEVERHGLI
QKRAADITKALVRMFKEMRDLILSETEQINIOQOASVYMQTLKROAKVDS
BQRNSTNVRIRDPTGOCIPRPIPKGMKRVKXVDDDEGWSGSKVKVKGRLTSSG
DGDLSVPFENETTRIRNEKDKRLMLEIFGPSTDDSDSGTEESFEMTSLMSLW
GKRKRSNRKTRHRIENHVGQHKYQDVRKCLYCETLVHSFSLIGVVSMMQMD
IQVLLSMDASVIFNLNPQVSCPKPNTLKSCEVCLILT"
20721..24395
/gene="F3H9.3"
join(20721..20817,21062..21222,21318..21365,21455..21514,
21709..21837,21988..22098,22183..22286,22430..22488,
22580..22674,22757..23426,23624..23816,23957..24048,
24148..24395)
/gene="F3H9.3"
/note="Unknown protein"
/codon_start=1
/protein_id="AAF98423.1"
/db_xref="GI:9795605"
/translation="MDVSKVSVFGRNFAVLVKVGPVDPDLSLNSSSGILLPRDIFL
SGEVAAKVLPFAGQDMALVLVAVSVVEPFLTGHRTSSISIQDPVKLI PGAMIIMVE
GQLKEAPEWPVPAQLLSLVDVPVSAALQSLIEASSGSKSDGWDICWSLVSAAAGS
QPSNIEHPSGLQDLPHPNANFMAKSTRMILVPLSLGOPSNNFASSSSKGDT
LVALGPSLSPVNFNSYPSGSLKKSLMADVRCLPMEGAPVEAKNGHLJGILIR
PLRQNSGVELQQLVVPKGAITACSHLLLEEPSVEGKASQWSEVLSVKSASIPAQ
VAIBKMSVCLITVNDGWSAGIILNHLGHLLTNAHLLEPWRKYKGSGVGEKFPYV
LGAEEFSTGSKFWQKSQTLPRKAPRNHYSSGENTREYKHNFLQTCRHRDVRVLC
LDSWTWPANTVYICKEOALIALQLEYVPGKLOPITANFSPPPLGTAAHVGHFLG
PRCLGPSICGVAKVAVKARLNTQSIQSOEAFPMLETTAAVHPGSGAVLNS
SGHMLGLVSNRAGCAGTVIPLHNFSPCAVLPIFKFAEDMONTTILQTLDDPSEEL
SIVNLMPSLSPKTEQSLPNLPLKLLKDNNGKQTKGSQFAKFIATQDMFVKPTKLSRD
VIPSKL"
31599..32534
/gene="F3H9.4"
31599..32534
/gene="F3H9.4"
/note="Hypothetical protein"
/codon_start=1
/protein_id="AAF98424.1"
/db_xref="GI:9795606"
/translation="MQSKNMIVASSHQOQQOQQPQPOPQLKPCRCDSNSTKFCYYNN

YLSQPRHFCKACKRYWTRGCTLRNVDPVGSYRKNKRVKRPSTATTTTASTVSTNSS
SPNNHOISHESSNNHPLFYGLSDHMSSCNNLPMIPSRFSDSSKTCSSSGSSEFL
SSGFSSLSALCLGLPHQMSHDHTINGSFINNSTNPKFLLSGLFGSSMSSSSTLQHP
HKPMNNGDWLGQSHLQTLASQLDHVGNEDMKYKEGKLDQISGNIFMSSSSSL
DPNYSNNMMNNSAVVNGAWLDPDTNNNVGSSLTSLI"
45230..48424
/gene="F3H9.5"
join(45230..45724,45882..45956,46051..46951,47136..47182,
47292..47365,47715..47984,48407..48424)
/gene="F3H9.5"
/note="Hypothetical protein"
/codon_start=1
/protein_id="AAF98425.1"
/db_xref="GI:9795607"
/translation="MDNLFPPSSNANSVOELSMDPNNNRSHFTTIVTYDHHQAPPH
FLPPFSYFVEQMAAVNMPQVYILSECYPIPTVTGSEFSLVGNPCWQERGFLOP
RMTKMARINRKNAMSRNNSPNSFSLVDSKQMLMLNLKNVQISDKQSYQDS
TFONKLRVLCEKELKNSDVGLGRIVLPKRDAEANLPKLSKREGI VVQWRDVFMSQS
WSFKYKFSWNKSRMYVLENTGFEVKONGAEIGDFLTYDESKNLYFAMNGNSKQN
EGNESRERNHYEEAMLDYIPRDEBEASIAMLIGNLNDHYPIPNLMDLTLTDLQHHQ
ATSSMPPEAHYVGSDDQVLDWL"
54225..55737
/gene="F3H9.6"
join(54225..54993,55427..55737)
/gene="F3H9.6"
/note="Unknown protein; proline rich amino acid sequence"
/codon_start=1
/protein_id="AAF98426.1"
/db_xref="GI:9795608"
/translation="MGPIGKSVLSVALMCFSTSSVFTEBVNHNKTOPSLAPAPAPYH
HGHHHPVPPHHHPHPHPHPKAPSVKAPVSPAPKPPVPKPPVPTKAPKPPV
PTKPPKPPVPPAKPEVPKPPVPTKAPKPPKPPKPPVPTKAPKPPKPPV
KPPVPTKAPKPPKPPKPPVPPVSPAKPPKPPVPTKAPKPPVSPPTKPPVTP
VPPPKPNSLVAVRGTVYCKSKYAAFNLLGAKPLEGATVKLVCKSKKNITARETT
DKNGYFLLLPKPTVNFPGRCRVLYLKDKYKCSKVSLFGGVDVGAPEKPLGKS

Query Match
Best Local Similarity 63.78; Pred. No. 5.8e-111;
Matches 682; Conservative 0; Mismatches 313; Indels 76; Gaps 1;

Qy 2 AGATGAAGATGAAGACAGTTCCTTGTATATCAATAACTGTATATCTTGGCCATTGGAACCT 61
Db 83626 AGACAAGATGGTGAAGGCTCTTGTGATCATAAACTGCATAATTCATAGCCATAGGAACCT 83685

Qy 62 GTGGAGGCGCTCTAATCATCGGCTCTACTTCCAAATGTTGCGGAAGGATCTGGTTTC 121
Db 83686 GTGGAGGCTCTTGAITATGCGTCTCTACTCAACAATGCGGTAAGGATTTGGTTCT 83745

Qy 122 CAAGCTTCCTTCAACCGTTGGTTGTCCACTCATATTTTCTCCCTCTTCTCTATCTTCTCC 181
Db 83746 CTAGCTTCTTGAACCTGACGGCTTCTCTGTATCTTATCTCTCTCTCTCTCTTACA 83805

Qy 182 TCCGCGCTGCTGCTGCTTGAAGAACCAAGAACGACTCCATTTTCTCATGAAACCTC 241
Db 83806 TTACCCGCGCAGAGAACAACAATGTGGGTGATAGTAGTCAAGTTCTTCTTATCAACCGC 83865

Qy 242 CTCTCTTATCGCGCTACTCGTTGTTGCTGCTGGGATTTGCAATTTACCTCTACT 301
Db 83866 GTCTCTTATCGCGCTGTTATTGTGAGGCATCTCTCAGGGTTTGATAACTACTTGTATG 83925

Qy 302 CTTACGGGTTAGCTTATATCCCTGTTTCTACTGCGGTCTTTGATCATCTCCCGCAATTAG 361
Db 83926 CATATGATAGCTTATCTCCAGTTTCTACAGCTGCTCTTATCATTTGCTTACAGTTAG 83985

Qy 362 GCITTCATGCTCTCTTTGCAATTTTATGTTGAAGCAAAAAGTTTCAACCTTTCTATAA 421
Db 83986 CTTTATAGCTATCTTCTCATTTCTTATGGTTAAACATAAGTTCACCTCTTTTACCATCA 84045

Qy 422 ACGCTATCGTTTGTCTCACTGGTGGTGGCTAGTCTTCCCTTAACCTGATAGTGACA 481
Db 84046 ATGCTGTGTGTGTGACTGTGTGGTGTGGTGTGGGAAATGATACCGGAACTGATA 84105

Qy 482 AGCTTGCAACGACAGACACACAAGGAATATGTTGGGTTTCATCATGACTCTTGGTGAG 541
```

Db	84106	
----	-------	--

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORSScience 301 (5631), 376-379 (2003)
22752273
12869764

2 (bases 1 to 1505)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iehi, Y., Ikeda, R., Imanura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurotsaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otsu, Y., Otsu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yabuchi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission

TITLE

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kamondai, Teikyo, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yabuchi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imanura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

source

TITLE

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oota, N., Ota, Y., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Takami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

Location/Qualifiers
 1. 1458
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="001-036-F06"

ORIGIN

Query Match 23.1%; Score 242; DB 8; Length 1458;
 Best Local Similarity 55.0%; Pred. No. 3.5e-61;
 Matches 529; Conservative 0; Mismatches 415; Indels 18; Gaps 2;

QY 20 TTCTTGTGTAATCAATAAAGTATATCTTGGCCATTTGGAACTGTGGAGGCGCTCTAATGA 79
 DB 202 TCCTCTGGCGCTCACTGCGGATGCTGTGTGGCCACCGGGCGGCGCATCCCTA 261

QY 80 TGCCTCTCTACTTCCAAATATGGGGGAAGATCTGGTTTCCAAGCTTCCTTCAAACCG 139
 DB 262 GCGCGCTCTACTTCAGCAAGGGGCGCACAGGAAGTGGCTCTGGCGTGGCTCGAGCCG 321

QY 140 TTGGTTGTCACTATTTCTTCCCTCTTCTCTTATCTTCTTCCTCGCGGTGCTGTGGC 199
 DB 322 CCGGGTGGCGCTGCTGTGTGGTGGCGTGTGGCGTCTTACCACAGCGCGCGCGCGCG 381

QY 200 TTGAAGAACAAAGAACGACTCCATTTTCTCTATGAAACCTCTCTTTATCGCGCTA 259
 DB 382 ACCGCGCG-----CGCCGCTGTCTCTACGCGCGGCGGCTGTGTGGCGGGCG 432

QY 260 TCGTTGTGTTGTGCTGGGATTTGACAAATACCTCTACTCTTACGGGTAGCTTATA 319
 DB 433 CGGTGCTCGCGCTCTCACCGCGCGGACGACTTGTCTAGCGCTAGGCGCTCGCTACC 492

QY 320 TCCTGTTCTACTGCTTGTGATCATCTCGCGCAATTAGCTTCTACTGCTCTCTTTTG 379
 DB 493 TGCCTGTGTCACCTCGCGCATCTCTATCTCCACGAGCTCGCGTTCACCGTCTTCTTCG 552

QY 380 CATTTTGTATGGTGAAGCAAAAGTTTCAACCTTTCATATAACGCTATCGTTTGTCTCA 439
 DB 553 CGTGCTCTATGTCGGCGAGCGGCTCACCGCGGAGCTGTAACCGCGTGGCGCTCTCA 612

QY 440 CTGGTGGTGGCTGCTGCTTGGCTTAACTCTGATAGTGAAGCTTGAACAGACAC 499
 DB 613 CTATCGCGCGCTGCTGCTGGGCTTCCACGGTCCAAAGGACCGACCGCGGGGTACCA 672

QY 500 ACAGGAATATGTTGTTGGTTGATCATGACTTCTGGTGAGCTTCTTCTATAGGGTTTA 559
 DB 673 CCGGAAGTACTGGAATGGGCTTCTCTACCTCGCGCGCGCGGCTGTACGGGCTCA 732

QY 560 TATTGCCACTTGTGAGCTTCTTACAAGAAATCT-----GFTCAGCGNATCACGT 610
 DB 733 TCCTGCGGCTGCTGAGCTGCGGTACAGACCGCGCGGCGGCGCGCGCTCACGT 792

QY 611 ATACGCTCGCGTTCGAGTTCACAGATGCTTATGCTTTGCTGCACTTGTGCTGCTCG 670
 DB 793 ACGCGCTGTGATGGAGATGACAGTGTCTATGGCTTCTTTCGCCACCGGCTCTGTACCG 852

QY 671 TGGGGATGTAGCCGCTGGGATTTCAAGGTGATACGAGGAAGCAAGAGATTTTAAAGC 730

DB 853 TCGTGTGTGATCGTCAACAAGGACTTCCAGCGCATCCGAGGGAGGAGGACGAGTACGAGC 912
 QY 731 TTGGAGAGTCTTTTGTACTATGTGGTATGTGTTCACGGCCATAAATCTGCAAGCATTTT 790
 DB 913 TGGTGAACGGGGTACTACGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 972

QY 791 TTGTGGAGCTATTGCGGTTGATCTTCTGTGCATCGTCTCTGTGCTCTTGGAAATTTATGTC 850
 DB 973 TCGTGGGCGCGCTCGGCGTCATCTTCTGCGTGCACACGCTGCTCGCGGATCATCATCG 1032

QY 851 GTGCTCTGCTTCGGTGAAGGATCTGGCGCGTATTTGCTTCCAGGAGAGATTTTCAGG 910
 DB 1033 CCGTGTTCATCCCGATCACGAGGTGCTGGGGGTCTATCTTCTCCACGAGAAATTCAGCA 1092

QY 911 CCGGGAAGGCTGCTCGCTTTGGCTCTCTCCCTCTGGGGATCAGTCTTATTTCTATGGAC 970
 DB 1093 GCGAAGAGGCGTGGCGTCTGCTCTGCGTCTGGGGCTCGCTCTCTACTCTCTACGGCG 1152

QY 971 AG 972
 DB 1153 AG 1154

RESULT 12
 AK064135
 LOCUS
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:002-102-G04, full insert sequence.
 ACCESSION AK064135
 VERSION AK064135.1 GI:32974153
 KEYWORDS FLI CDNA; oligo-capping.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1
 AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niihara, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oota, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.

TITLE japonica rice
 JOURNAL Science 301 (5631), 376-379 (2003)
 MEDLINE 22752273
 PUBMED 12869764

REFERENCE 2 (bases 1 to 1363)
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Itoh, Y., Ikeda, R., Imamura, K., Kagawa, I., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kawai, J., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Murakami, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niihara, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Ooka, M., Ooka, H.,

Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sato, Y., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Rakeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007]

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Sato, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oso, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Rakeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

source

1..1363
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="002-102-G04"

ORIGIN

Query Match 21.5%; Score 226; DB 8; Length 1363;
Best Local Similarity 53.7%; Pred. No. 2.3e-56;
Matches 536; Conservative 0; Mismatches 430; Indels 32; Gaps 2;

QY 9 GATGAAGACAGTCTTGTAATCATAACTGTATATCTTGGCCATTGGAACTGTGGAGG 68
DB 173 GTTCAGAGCCGCTCTCTGCTGCTCAACTCTGCTCAATGTTGGCTCGGCTGGCG 232
QY 69 CCTCTAATGATGCGTCTCTACTTCCAAAATGGTGGCGAAAGGATCTGGTTTCCAAGCTT 128
DB 233 GCGGCTCTCTCGCGGCTACTTCTCCGGGGGCAACCGCAAGTGGCTCTCCAGGCT 292
QY 129 CTTTCAAAACCGTGTGTTGCCACTATTTTCTTCCTCTTCTATATCTTCTTCCTCGGCG 188
DB 293 TCTTCAGACCGCGGCTGCGGCTGCTGCTGCTGCGGCGCTGCTGCTTCTCGTACTCTCAGC 352
QY 189 TCGTCTGTTGCTTGAAGAACAGAA-----CGACTCCATT 224
DB 353 CGCGCGCGCGGAGTTGAGACGACGGCTGGCGCTGGCGGCGGCGGCGGCGGCT 412
QY 225 TTTCCTCATGAACCTCTCTCTTTTATCGCGCTATCGTTGGTTGGTTCGTGGGATT 284
|||||

DB 413 GTTCTCATGACGCCACGCGCTCTGTGGCGTCCGCGCTCGTCGGGCTCATGACCGGCGT 472
QY 285 TGACAAATTACCTCTACTCTTACGGGTTAGCTTATATCCCTGTCTTCTACTCGGCTCTTGAT 344
DB 473 CGACGACCTCTCTACGCTTACGCGCTGGCGTACCTCCCGGTGTCCACCTCTCCATCCT 532
QY 345 CATCTCGCGCAATTAGCTTTCAGTCTCTCTTTGCAATTTTATGTGTGAAGCAAAAGTT 404
DB 533 CATCTCCAGCAGCTGGCTTTCAGCGGGGCTTCCGCTGCTGCTCGTCCGCGGCGGTT 592
QY 405 CACACCTTTTCAATAAACGCTATCGTTTGTCTCACTGGTGTGCGGCTAGTCTTGGCCT 464
DB 593 CACGCGGTTCTCGGTGAAGCGCGTCTGCTCAGCGTGGCGGCGGATGCTGGGAT 652
QY 465 TAACTCTGATAGTGAAGCTTGCAAAAGACGACACAAAGGAATATGTTGTGGTTTCAT 524
DB 653 GAACCGCGGGGACCGCGCGGGGGTGTCCGCGCGCAGTACTGCGCGGGTTGCG 712
QY 525 CATGACTTCTGGTGCAGCTCTTCTATAGGTTTATATGGC-----ACTTGTGAG 576
DB 713 CATGACGCTGCGCGCGGGGCTGTACGGGCTGCTGCTCCCTCATGGAGCTCAGCAG 772
QY 577 CTTTCTTACAGAAATCTGCTCAGCGAATCAAGTATACGCTCGCGCTCGAGTTCCAGATG 636
DB 773 GCGCACCGCGCGCGCGCGCGCTCAGGTACAGCTCGTCTGTCATGGAGATGCAGCTC 832
QY 637 GTCTTATGCTTGTGCTCACTTGTGCTGCTGCTGGGATGCTAGCGGCTGGCGATTTTC 696
DB 833 GTCATCGGCTTCTGTCGCGCCAGCGCTTTCAGCGCGCTCGGCATGCTCGTCAACAGATTC 892
QY 697 AAGGTGATAGCAGGAGAAAGAGATTTTAAAGCTTGGAGAGTCTTTGTACTATGTGTG 756
DB 893 CACGCAATCCAGAGAGAGCCCATGAATTCGGTCTGGGCAAGGGGCTACTACTGCTC 952
QY 757 ATTGTGTTACGCGCAATACTTGGCAAGCAATTTTGTGGAGCTATTTGGGTTGATCTTC 816
DB 953 CTGGCGGATCGGCTGCCATGTACAGTGTCTTCTCGGACGATCGGCGCATCTTC 1012
QY 817 TGTGATCGTCTGCTGCTCTGGAATATGTTAGTCTGCTCTGCTTCCGCTGACCGTATC 876
DB 1013 TACGGCTCGGCGTGTCTGCGGCGCTCATCATGACCGTGTCTCATCCCGGTACCGAGGTG 1072
QY 877 TTGGCGCTCATTTGCTTCCAGGAGAGTTTTCAGCGCGGAAAGGTGTGCTTTTGGCTCTC 936
DB 1073 CTGCGCGCTATGTTCTTCCAGAGCAATTCAGCGGCAAGGGGCTGCGCTCGCTCTC 1132
QY 937 TCCCTCTGGGATCAGTCTCTTATTTCTATGGACAGGT 974
DB 1133 TCGCTATGGGCTTCGCTCTCTACTTCTACGGCGAGGT 1170

RESULT 13

AY096558

LOCUS

DEFINITION Arabidopsis thaliana putative protein (At4g18200) mRNA, complete cds.

ACCESSION

VERSION AY096558.1

KEYWORDS FLI CDNA.

SOURCE

ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1165)

REFERENCE

AUTHORS

Yamada, K., Ban, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE JOURNAL REFERENCE AUTHORS

Arabidopsis Open Reading Frame (ORF) Clones

Unpublished
2 (bases 1 to 1165)
Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN
Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

TITLE JOURNAL COMMENT

The Salk, Stanford, PGENC (SSP) Consortium members constructed and
sequenced the pUNI (ORF) clones using the RAPL cDNAs: Yamada,K.,
Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGENC) and Seki, M. (RIKEN GSC) contributed equally
to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
/PGENC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.

FEATURES source

Location/Qualifiers
1. .1165
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U18528"
/ecotype="Columbia"
/notes="This clone is in pUNI 51."
1. .1165
/gene="At4g18200"
1. .1134
/genes="At4g18200"
/codon_start=1
/evidence=experimental
/product="putative protein"
/protein_id="AAM20208.1"
/db_xref="GI:20465497"
/translation="MGISQVHYCNGDQNLLEANLLDHEETSFSPQTKNCKRWLRVSI
YAFVFCPLATLGRLYENGKSTVYVLLQLGFPVLILPFRFRIROKPSKTDI
NFSQSPSTFLASVILCTGLLVSAVAYLSAVGLLPLVFSLLILASQLAFAPFSYF
LNSOKFTPLVNSLFLTVSSALLVNTDSENTNVSQVYVIGFICTIGASAGIGLV
LSLIOLLFRKVTQHTSSAVLDLANVQSLVATCVLIGLFPASGEWRTLPSEMNKLG
KYSYLITLASAIFMQVYVGVCGVLIPSSSVFSNSITVAIGLPVIVVAIVFHDKMD
ASKIFSIILAIAGFLSFVYQHYLDLDEKKLKTCTQKPVBEETQL"
1135. .1165
/genes="At4g18200"

gene

CDS

3' UTR

ORIGIN

Query Match 16.0%; Score 167.6; DB 8; Length 1165;

Best Local Similarity 48.6%; Pred. No. 9e-39;

Matches 492; Conservative 0; Mismatches 514; Indels 6; Gaps 1;

QY 1 AGAGTGAAGATGAAGACAGTCTTGTAAATCATAAACTGTATATCTTGGCCATTGGAAAC 60

Db 100 AGAACTGTGAAGGTGGCTCGGTGCTCATATACGCAATCTTGTCTATCTTCTGCCAA 159

QY 61 TGTGGAGGCCCTCTAATGATGCGTCTCTACTTCCAAAATGGTGGCGAAAGGATCTGGTTT 120

Db 160 CCACCTTGCTACAGTCTCTGGTAGACTGTAATAATGGAGGAAACACATATGTG 219
QY 121 CCAAGCTTCCTTCAAAACCGTGGTGTGCCACTCATTTCTTCCCTCTCTCTTATCTTTC 180
Db 220 GTAACACTCTTCAACTCAATGGCTTCCTGTACTGATCTCTTCGGCTTCTTTTCTCGA 279
QY 181 CTCGCCCGTGTGCTTGGCTTTGAAGAAACAAGAAACAGACTCCATTTTTTCTCATGAACCT 240
Db 280 ATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCCCTTCCCTTCAACCCCTT 339
QY 241 CCTCTCTTTATCGCGCTATCGTTGTTGCTGCTGGGATTTGACAAATTTACCTCTAC 300
Db 340 GCATCGGTTTAC-----TTGTGCACCTGGACTGCTAGTGTCTGTATATGCTTTATTGTCT 393
QY 301 TCTTACGGGTTAGCTTATATCCTGTTCTACTGCTGCTTTGATCATCTCTCGCGGAATTA 360
Db 394 GCAGTTGGTGTCTTACTTACCAGTCTCTACTTTCTCCTCATCTTGGCCTCACAGTTG 453
QY 361 GGCCTTCACTGCTCTCTTTGCAATTTTTATGTTGAAGAAAGTTTCAACACTTTCACTATA 420
Db 454 GCCTTCACTGCCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTCTCTTGTAGTGC 513
QY 421 AAGCTATCGTTTGTCTCACTGCTGGTGGCGTAGTCTTTCCTTAACTCTGATAGTAGAC 480
Db 514 AATTCTTTGTTTCTCCTTACGGTTCCTCTGCCCCTCCTCGTGGTCAACACTGATTCAGAA 573
QY 481 AAGCTTGCAGAACGACACACAAAGAAATATGTTGTTGGTTCATCATGACTCTTGTGTGA 540
Db 574 AACACAAATATGATCTAGAGTAAATGATGATGGGTTTCACTGTACCAATGTTGTGT 633
QY 541 GCTCTTCTCTATGGGTTTATATTTGCCACTTTGTCAGACTTTCTTCAAGAAATCTGCTCAG 600
Db 634 TCCGCTGGGATTTGGATTTGTAATACTCTCTGATACAACTGCTCTTCAGGAAAGTTTTCAG 693
QY 601 CGAATCAGTATACGCTCGGCTCGAGTTCCAGATTCGATGCTTATGCTTGTCTGCCACTTGT 660
Db 694 AAGCATACATCCTCAGCAGTCTCTGGCAATTACCACTCTCTAGTTGCAACTTGT 753
QY 661 GTCTGCTCGTGGGATGCTAGCCGCTGCGCAATTTCAAGGTGATAGCAGGAGAAACAAGA 720
Db 754 GTGGTACTACTAGGACTGTTTGCAAGTGGAGAGTGAGAACTCTGCCAAGTAGATGAGA 813
QY 721 GATTTTAAAGCTTGGAGAGTCTTTGTAATATGTTGATGTTGTTTTCACGGCCATAATCTGG 780
Db 814 AACTACAACTGGGAAAGTGTATATATCTTGACTTTGGCCTCAGCAGTATTTTCTGG 873
QY 781 CAAGCAATTTTGTGGGAGCTATTGGGTTGATCTTCTGTGCATCGTCTCTGTCTCTCGGA 840
Db 874 CAAGTATACACTGTTGTTGTTGGGATTAATCTTCGAGTCTTCTTCTGTGTCTTCCAAT 933
QY 841 ATTATGCTCAGTCTCTGCTTCGGTGGAGTCTTGGCGGTGATCTTGGCGGTGATTTGCCAGG 900
Db 934 TCCATAACTGCTGTGGAGTGCCTATAGTTCAGTTGTAGCAGTATAGTTTTCATGAT 993
QY 901 AAGTTTCAAGCGGGGAAAGGTGTCTGCTTTGGCTCTCTCCCTCTGGGGATCAGTCTCTTAT 960
Db 994 AGATGGATGCATCAAGATTTTCTCCATCATTTTAGCTATCTGGGCTTCTCTTCAATTT 1053
QY 961 TTCTATGACAGGTTTAAATCCGAGGAGAAAGTAAAGGCTCAGGATACACAAC 1012
Db 1054 GTCTATCAGCACTACCTCGACGAAAGAAAGTTGAAGACTTGGCAGACAAAAC 1105

RESULT 14

AY074546

LOCUS

DEFINITION

Arabidopsis thaliana unknown protein (At4g18200) mRNA, complete cds.

ACCESSION

AY074546

VERSION

AY074546.1

KEYWORDS

FLI CDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1 (bases 1 to 1428)

AUTHORS

Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

TITLE

Arabidopsis Full Length cDNA Clones

REFERENCE

2 (bases 1 to 1428)

AUTHORS

Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

TITLE

Direct Submission

JOURNAL

Submitted (16-JAN-2002) Plant Gene Expression Center, 800 Buchanan

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA': Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

source

1..1428
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="4"
 /clone="RAFL09-36-N15 (R18628)"
 /ecotype="Columbia"
 /note="This clone is in a modified pBluescript vector (FLC-1) as a BamHI/XhoI insert."

gene

1..1428

5'UTR

1..71

CDS

72..1205

/gene="At4g18200"

/codon_start=1

/evidence=experimental

/product="unknown protein"

/protein_id="AAL69512.1"

/db_xref="GI:1849121"

/translation="MGISQVHCNGDONLEANDHRETESFSVPQTKNCKRLRVSI
 YAIWFCQPLATVGLRYXGCKSVYVTLQIGFVLLIRFRPSRIRQPSSTDT
 NFSQSPSETTLASVYLCGLVSAIYLSAVGLYLPVSTFSLIASQLAFSTF
 LNSQKFTPLVNSLFLVTSALLVNDSTNVRQVQVIGICTIGASAGIGLV
 LSLIQLFKRVFTKHTSSAVLDLANYQSLVATCVVLIGLFASGEWRTLPSEMRNYKLG

KVSYILTASAAAIWFQVYTVGCVGLIFESSVSFNSITAVGLPIVPPVAVIVFHKMD
 ASKTFSIILAIWGLSPVYQHYLDEKXKTKQTKPVEEIQTL"
 1206..1428
 /gene="At4g18200"
 misc_difference 1412
 /gene="At4g18200"
 /note="not present in genomic sequence"

ORIGIN

Query Match 16.0%; Score 167.6; DB 8; Length 1428;
 Best Local Similarity 48.6%; Pred. No. 9.3e-39;
 Matches 492; Conservative 0; Mismatches 514; Indels 6; Gaps 1;
 QY 1 AAGATGAAGATGAAGACAGATTCTTGTAAATCATAACTGTATATTCTTGGCCATTTGGAAC 60
 DB 171 AAGAACTGTAAAGGTGGTCCGTGTCTCCATATACGAATCTTTGTGATCTTCTGCCAA 230
 QY 61 TGTGGAGCCCTCTAATGATGCGTCTCTACTTCCAAAATGGTGGCGAAAGATCTGGTTT 120
 DB 231 CCACCTTGCTACAGTTCTTGGGTAGACTGTACTATGAAAATGGAGGGAAGACATATGTG 290
 QY 121 CCAAGCTTCCCTTCAACCGTTGGTGTGCGACTCAATTTTCTCCCTCTTCTCTTATCTTTC 180
 DB 291 GTAAACACTTCTTCAACTCAITGGCTTCCCTGTACTGATCTCTTCCGCTTCTTTCTCGA 350
 QY 181 CTCCGCCGTGCTGCTTCCCTTGAAGAAACAAGAACGACTCCATTTTCTCTCATGAACCT 240
 DB 351 ATCAGGAACCCAAATCAACAGATACAAATTCAGTCAGTCCCCTTCTTCCACACCCCTT 410
 QY 241 CCTCTCTTATTCGCCGCTGCTTGTGGTTGCTCGTGGGATTTGACAAATACCTCTTAC 300
 DB 411 GCATCGGTTTAC-----TTGTGCACCTGGACTGTCTGTCTTATGCTTATTTGTCT 464
 QY 301 TCATTACGGTTAGCTTATATCCCTGTTTCTACTCGCTTGTGATCATCTCCGCCAATTA 360
 DB 465 GCAGTTGGGTGCTTTTACTTACCACTCTCTACTTTCTCCCTCATCTTGGGCTCAGTTG 524
 QY 361 GGCTTCACTGCTCTCTTGGCAATTTTATGGTGAAGCAAAAGTTTCAACCTTTTCACTATA 420
 DB 525 GCCTTCACTGCTTTTCTCATATTCTTAACTCGGCAAAAGTTCACTCTCTTTGATAGTC 584
 QY 421 AACGCTATCGTTTGTGCTCAGTGGTGGTGGCTAGTCTTCCCTTAACTCTGATAGTAGC 480
 DB 585 AATCTTTGTTTCTCTTACGGTTTCTCTGCCCTCTCTCGTGGTCAACACTGATTCAGAA 644
 QY 481 AAGCTTCAACAGACACACAGGAATATGTTGTTGGGTTTCATCATGCTCTTGGTGCA 540
 DB 645 AACACAAACAAATGATCTAGAGTACAATATGTGATTTGGGTTTCATCTGTACCATTTGGT 704
 QY 541 GCTCTTCTATGGTTTATATTGCCACTTGTGAGCTTTTCTTACAAGAAATCTGGTCAG 600
 DB 705 TCCGCTGGGATTGGATTGGTACTATCTCTGATACAACTGCTCTTCAGAAAGTTTTCACG 764
 QY 601 CGAATACGATACGCTCGCGCTCGAGTTCAGATGGTCTTATGCTTTTGTGCGCACTTGT 660
 DB 765 AAGCATACATCCTCAGCAGTCTGGAAGTGGCAATTAACAGTCTCTAGTTGCAACTTGT 824
 QY 661 GTCTGCCCTCGTGGGGATGCTAGCCGCTGGCGATTTCAGGTGATAGCAGGAGAGCAGA 720
 DB 825 GTGGTACTCATAGGACTGTTTGAAGTGGAGAGTGGAGAACTCTGCCAAGTGAAGATGAGA 884
 QY 721 GATTTTAAAGCTTGGAGAGTCTTTGTACTATGTGGTGAATGTTTTCAGGGCCATAATCTGG 780
 DB 885 AACTACAACTGGGGAAGTGTCAATATCTTGACTTTGGCTCAGCAGCATTTTCTGG 944
 QY 781 CAAGCAATTTTGTGGAGCTATTGGGTTCATCTTCTGTGCACTGCTCTGGTCTCTCGGA 840
 DB 945 CAAGTATACACTGTTGGTGTGGGATTAATCTTCAGTCTTCTTCTGTGTTCTCCAAT 1004
 QY 841 ATTATGGTCACTGCTCTGCTTCCGGTCAACGGTGAATCTTGGCCGTCATTTGCTTCCAGG 900
 DB 1005 TCCATAACTGCTGTGGGACTGCCTATAGTTCCTAGTTTGTAGCAGTGATAGTTTCCATGAT 1064

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:44:43 ; Search time 4791.13 Seconds
(without alignments)
10669.751 Million cell updates/sec

Title: US-09-913-767-6
Perfect score: 1081
Sequence: 1 caaatccacagttcaagat.....ataagctgatattttgaaag 1081

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1077.8	99.7	1081	6	AX033549 Sequence
2	1043.8	96.6	1047	6	AX412868 Sequence
3	1008.8	93.3	82697	8	ATT9A21
4	1008.8	93.3	194143	8	ATCHRIV48
5	974.8	90.2	1106	8	AF370622 Arabidops
6	835.8	77.3	1071	6	AX033550 Sequence
7	406.8	37.6	1293	6	AX033547 Sequence
8	406.8	37.6	58427	8	ATF15U5
9	406.2	37.6	3387	6	AX507517 Sequence
10	406.2	37.6	3387	6	AX652015 Sequence
11	385.2	35.6	73179	8	AP006383 Lotus cor
12	369.6	34.2	1194	6	AX033348 Sequence
13	368.2	34.1	1165	8	AY096558 Arabidops
14	368.2	34.1	1428	8	AY074546 Arabidops
15	307	28.4	1476	8	AK071254 Oryza sat
16	307	28.4	1617	8	AK072660 Oryza sat
17	307	28.4	175645	8	AP004096 Oryza sat
18	298.8	27.6	1251	8	BT003358 Arabidops
19	298.8	27.6	1372	8	AY062782 Arabidops

20	298.8	27.6	115421	8	T12C22	AC020576 Sequence
21	298.4	27.6	1145	6	AX033546	AX033546 Sequence
22	294	27.2	110036	8	AC142095	AC142095 Medicago
23	262	24.2	133299	8	OSJN00098	AL606658 Oryza sat
24	262	24.2	145739	8	OSJN00157	AL662957 Oryza sat
25	258.8	23.9	1486	8	AK119483	AK119483 Oryza sat
26	234.2	21.7	83698	6	AX540653	AX540653 Sequence
27	234.2	21.7	83698	8	AB010072	AB010072 Arabidops
28	225.8	20.9	1117	8	AY096468	AY096468 Arabidops
29	225.8	20.9	1333	8	AY074535	AY074535 Arabidops
30	225.8	20.9	1403	8	AY087844	AY087844 Arabidops
31	225.8	20.9	104386	8	AT132A17	AL161813 Arabidops
32	225.8	20.9	199362	8	ATCHRIV24	AL161512 Arabidops
33	222.8	20.6	2175	6	AX654201	AX654201 Sequence
34	211.6	19.6	1152	6	AX653666	AX653666 Sequence
35	208.2	19.3	721	11	BX664674	BX664674 Arabidops
36	205	19.0	1279	8	BT014359	BT014359 Lycopersi
37	186.2	17.2	733	8	AF083706	AF083706 Arabidops
38	180.4	16.7	1361	8	AK071680	AK071680 Oryza sat
39	145.6	13.5	1049	6	AX033545	AX033545 Sequence
40	138.2	12.8	1067	6	AX033553	AX033553 Sequence
41	136	12.6	1225	6	AX033544	AX033544 Sequence
42	136	12.6	1227	8	AF078531	AF078531 Arabidops
43	134.4	12.4	1102	8	BT005504	BT005504 Arabidops
44	134.4	12.4	1250	8	AK117664	AK117664 Arabidops
45	129.2	12.0	1155	6	AX653139	AX653139 Sequence

ALIGNMENTS

RESULT 1
AX033549
LOCUS AX033549 1081 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 6 from Patent WO049152.
ACCESSION AX033549
VERSION AX033549.1 GI:10280293
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsais.
REFERENCE 1
AUTHORS Andre B., Buerkle, L., Frommer, W.B. and Gillissen, B.
TITLE Nucleic acids that code for a nucleobase transporter
JOURNAL Patent: WO 0049152-A 6 24-AUG-2000;
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLF B (DE) ;
GILLISSEN BERND (DE)

FEATURES
source
location/Qualifiers
1..1081
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN	Query Match	99.7%	Score 1077.8;	DB 6;	Length 1081;
	Best Local Similarity	99.8%;	Pred No. 1.3e-253;		
	Matches 1079;	Conservative 0;	Mismatches 25;	Indels 0;	Gaps 0;
Qy	1	CAAAATCCAAAGTTC	CAAGATGAAAGAAATTCAGTCAGTAGCAGCGCAAGCAAGTATCTC	60	
Db	1	CAAAATCCAAAGTTC	CAAGATGAAAGAAATTCAGTCAGTAGCAGCGCAAGCAAGTATCTC	60	
Qy	61	ACTCTAACACATACA	ACGGTGGCTCAGGAGTACTATGACACCGGAGGAAACAGTAAA	120	
Db	61	ACTCTAACACATACA	ACGGTGGCTCAGGAGTACTATGACACCGGAGGAAACAGTAAA	120	
Qy	121	TGGCTAGCAACGGT	AGTTCACTTGTGGCTTTCTCTGCTACTTCCATATATATCTTG	180	
Db	121	TGGCTAGCAACGGT	AGTTCACTTGTGGCTTTCTCTGCTACTTCCATATATATCTTG	180	
Qy	181	TCATTTAAACACAT	GCACAACTGATAGATGAAAAAGAAACCTCACCTAGGAACCGT	240	

QY 902 ATGGGTTAAAGTGATTCTATGATCCTAGTATTGGGGTTTCACTTCTATGCTTACC 961
 |||||
 Db 16226 ATGGGTTAAAGTGATTCTATGATCCTAGTATTGGGGTTTCACTTCTATGCTTACC 16285
 |||||
 QY 962 AACAAATATCTTGATGACAAAACCTTGAAGAAAATCATGAATCACAACAAGATCCC 1021
 |||||
 Db 16286 AACAAATATCTTGATGACAAAACCTTGAAGAAAATCATGAATCACAACAAGATCCC 16345
 |||||
 QY 1022 CTGACCACCAAGAGAGAGTCAACTTGGCAATCAAAATAAGCTGATATTTTGAAG 1081
 |||||
 Db 16346 CTGACCACCAAGAGAGAGTCAACTTGGCAATCAAAATAAGCTGATATTTTGAAG 16405
 |||||

RESULT 4

ATCHRIV48

LOCUS ATCHRIV48 194143 bp DNA linear PLN 16-MAR-2000
 DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 48.
 ACCESSION AL161548

VERSION AL161548.2 GI:7268604

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 111084)
 Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and

AUTHORS

Mayer,K.F.X.

JOURNAL

Unpublished

REFERENCE

2 (bases 107966 to 194143)
 Hilbert,H., Braun,M., Holzer,E., Brandt,A., Duesterhoeft,A.,

AUTHORS

Mewes,H.W., Lemcke,K. and Mayer,K.F.X.

JOURNAL

Unpublished

REFERENCE

3 (bases 1 to 194143)
 Direct Submission

AUTHORS

EU Arabidopsis sequencing, project.

TITLE

Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer

Biochemie, Am Klopferspitze 18a, D-82152 Martinsried, FRG, E-mail:

lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project

Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge

Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,

E-mail: michael.bevan@bbsrc.ac.uk

Information on performance of analysis and a more detailed

annotation of this entry and other sequences of chromosomes 3, 4

and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/tha/>

this fragment has an overlap with ATCHRIV47 at the 5' end and an

overlap with ATCHRIV49 at the 3' end.

Location/Qualifiers

1. .194143

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/variety="Columbia"

/db_xref="taxon:3702"

/chromosome="4"

6377..9293

/gene="AT4g18120"

join(6377..6796,7122..7238,7347..7501,7578..8456,

8543..8627,8716..8784,8871..9293)

/note="similarity to Me12-like protein, Arabidopsis

thaliana, gb:D86122

contains EST gb:W43784, A1998063.1"

/codon_start=1

/product="putative protein"

/protein_id="CAB78814.1"

/db_xref="GI:7268605"

/db_xref="GOA:Q9SVV9"

/db_xref="TrEMBL:Q9SVV9"

/translation="WASGLDFSGGIGNWLDGDSHPICNMLPDDEELFSLMDLNL

LSSLATLDLEDYDLFSGGGLELETDPYSLNKGFRMGFADSNVNPQIFQN

GVGSAGHPHYGEHPSTRILFVRNINSNDESELQALFDNPFSEKDNVGLVFNLA

SVSNRDLNIFGVYGEIKREITETPNKRHHKFEVDFVRSADAALKALNRTEIARI

KLEHSRPGCARRRNMMLQMNPELEQDDSYSLNHVESPLASSPIGHWNRNSPIDHPLQSF
 QPQFGLNSLPTKNIRYDFEFMSKTVANNDQGRFRFSLDHLFSSSYNNASHKASTF
 QOQFQSGVSGFLGSHPSHVETLSGSEFLWGSFSSAWPVNPTSSNREHFPYSA
 QNGSLQHLHIGSAPSGFFPRSPETSSMGSAFRCASGNMNAORNLRETSSNFKMLS
 APRRSQTLFTNGSYLWPAATWVSIDDPEDSGNQFPDSNGNQADIKIQFQLDSKIMR
 GEDPRTTUMIKNIPNKYTRNMLLAIDAIDKNSGTDFLYLPIDFKNKNVGYAFINNV
 PKFTIALYEAFNGKKDKFSEKSEKASLAYARIQGAALIAHFQNSLNMDRRCCQIV
 FDGSSKYFVNVLQOQTHDKMIYLFQVILNMWLTHTCLFETRSGSGBFAARASNSTG
 FLIFLSARTHIKADSQNVNHTASGLKRVKQVFLG"

6377..6796

/gene="AT4g18120"

/number=1

6797..7121

/gene="AT4g18120"

/number=1

7122..7238

/gene="AT4g18120"

/number=2

7239..7346

/gene="AT4g18120"

/number=2

7347..7501

/gene="AT4g18120"

/number=3

7502..7577

/gene="AT4g18120"

/number=3

7578..8456

/gene="AT4g18120"

/number=4

8457..8542

/gene="AT4g18120"

/number=4

8543..8627

/gene="AT4g18120"

/number=5

8628..8715

/gene="AT4g18120"

/number=5

8716..8784

/gene="AT4g18120"

/number=6

8785..8870

/gene="AT4g18120"

/number=6

8871..9293

/gene="AT4g18120"

/number=7

9594..13230

/gene="AT4g18130"

complement(join(9594..9837,9944..10237,10323..11130,

11238..13230))

/gene="AT4g18130"

complement(join(9594..9837,9944..10237,10323..11130,

11238..13230))

/gene="AT4g18130"

/note="Contains ATP/GTP-binding site motif A (P-loop)

AA634-641;Phytochrome chromophore attachment site

AA317-326

contains EST gb:T21020, AA395462"

/codon_start=1

/product="phytochrome E"

/protein_id="CAB78815.1"

/db_xref="GI:7268606"

/db_xref="GOA:P42498"

/db_xref="Swiss-Prot:P42498"

/translation="MGFESSSAASNMKPPQKNTAQYSDAALFADFAQSIYTGKS

FNYKSVISPPNHVPDHSHTAYLGNIQRGVLQVPGCLIAVEEFSFRILGSDNSDF

LGLLSTSHSGEFDKVGKIGIDARTFTFTSSGASLSKAASFTISLLNPVLVHSR

TTQKPFYALLHRIDAGIVMDLEPAKSGDPALTLAGAVSQSKLAVRAISRSLPGGDI

GALCDTVVEDVQRUTGVRVNVYQFHEDDHGEVVSIEIRSDLEPLGLHUYPATDIPQA

ARFLFQNRNVRMI CDCNATPVKVQVSELSKPLCLVNSTLRPHGCHTQYMANMGSA

SLALAIYVVKGDSSKLMGLVVGHCSPRYVFPFLRYACEFLMQAGLQLOMELQASQ

QY 1022 CTGACCCACGAGAGAGAGTCAACTTGGCAATCAAAATAGCTGATATTTGAAAG 1081
|||||
Db 44733 CTGACCCACGAGAGAGAGTCAACTTGGCAATCAAAATAGCTGATATTTGAAAG 44792
|||||

RESULT 5
AF370622 1106 bp mRNA linear PLN 30-APR-2001
LOCUS Arabidopsis thaliana putative protein (T9A21.60) mRNA, complete
DEFINITION cds.

ACCESSION AF370622
VERSION AF370622.1 GI:13877726
KEYWORDS FLJ_CDNA.

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1106)
AUTHORS Lam,B., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R.,
Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H.,
Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission

TITLE Submitted (17-APR-2001) DNA Sequencing and Technology Center,
JOURNAL Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

COMMENT e-mail for correspondence: arab@sequence.stanford.edu
This clone was isolated by RT-PCR.

FEATURES
source
1. .1106
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="IV"
/clone="C531"
/ecotype="Columbia"
1. .1106
/gene="T9A21.60"
657. .1106
/gene="T9A21.60"
/codon_start=1
/product="putative protein"
/protein_id="AAK43941.1"
/db_xref="GI:13877726"
/translation="MDMIYVSLVASCVSVGLFASSEWKTLSSEMDNYKHGKVSIM
NLWTATVQVFSIGTGLFELSLFSNAISVLGLPVVPILAVIIFHDKMGLKVIS
MLAIWGFSTVYVQYLDLDDKLNKHEITTTESPPPEAEESTWQSK"

ORIGIN

Query Match 90.2%; Score 974.8; DB 8; Length 1106;
Best Local Similarity 99.8%; Pred. No. 2.4e-228;
Matches 975; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 89 GAGTATCTATGACACGAGGAGAAACAGTAATGGCTAGCAACGGTAGTTCAACTGTGTG 148
|||||
Db 130 GAGTATCTATGACACGAGGAGAAACAGTAATGGCTAGCAACGGTAGTTCAACTGTGTG 189
|||||
QY 149 GCTTTCTGCTGCTTCCATATATATCTTGTCAATTTAAACACATGCAACACTGATA 208
|||||
Db 190 GCTTTCTGCTGCTTCCATATATATCTTGTCAATTTAAACACATGCAACACTGATA 249
|||||
QY 209 GAGTGGAAAGAACCTCACCTAGGAACCGTGTATTGGTTTACGTAGTCTGGACTTC 268
|||||
Db 250 GAGTGGAAAGAACCTCACCTAGGAACCGTGTATTGGTTTACGTAGTCTGGACTTC 309
|||||
QY 269 TTGTAGGACGAGATTCTCTATCTGTACTCCATTGGACTTCTTTACTTACCGTTTCTACCT 328
|||||
Db 310 TTGTAGGACGAGATTCTCTATCTGTACTCCATTGGACTTCTTTACTTACCGTTTCTACCT 369
|||||
QY 329 ATTCCCTGATCTGTGCATCTCAGTTAGCCTTCAATGCTTCTCTCTTTTCTTAATCT 388
|||||

Db 370 ATTCCCTGATCTGTGCATCTCAGTTAGCCTTCAATGCTTCTCTCTTTTCTTAATCT 429
|||||
QY 389 CACAAAACTTACCCCTATCATTTTAAATTTCTCTTTTCTCTTAACTATATCTTCCACCC 448
|||||
Db 430 CACAAAACTTACCCCTATCATTTTAAATTTCTCTTTTCTCTTAACTATATCTTCCACCC 489
|||||
QY 449 TACTTGCATTCAATTAATGAGGAGACAGACTCCACAAAAGTTTACAAAAGGAGAGTATGTCA 508
|||||
Db 490 TACTTGCATTCAATTAATGAGGAGACAGACTCCACAAAAGTTTACAAAAGGAGAGTATGTCA 549
|||||
QY 509 AAGGTTTCATATGACACCGTGTCTGCGTCTGCTGGTGTATGGTCTAGTCTTATCCCTACAAC 568
|||||
Db 550 AAGGTTTCATATGACACCGTGTCTGCGTCTGCTGGTGTATGGTCTAGTCTTATCCCTACAAC 609
|||||
QY 569 AGCTAGCCCTTTCTAAAGTCTTAAAGAACCAAAATTTCTCAGAAGTTTATGGATATGATAA 628
|||||
Db 610 AGCTAGCCCTTTCTAAAGTCTTAAAGAACCAAAATTTCTCAGAAGTTTATGGATATGATAA 669
|||||
QY 629 TCTACGTGAGTCTAGTGGCCAGTTGTGTAGCGTGGTGGGGCTTTTGTAGCAGTGAAT 688
|||||
Db 670 TCTACGTGAGTCTAGTGGCCAGTTGTGTAGCGTGGTGGGGCTTTTGTAGCAGTGAAT 729
|||||
QY 689 GGAAAACTTTGAGCAGTGAATGGATTAATACTACAAACATGGGAAGGTATCTTACATTATGA 748
|||||
Db 730 GGAAAACTTTGAGCAGTGAATGGATTAATACTACAAACATGGGAAGGTATCTTACATTATGA 789
|||||
QY 749 ACCTAGTGTGGACAGCTGTACCTGCGAGTTATCTCCATCGTGGCAGCAGGACTGATCT 808
|||||
Db 790 ACCTAGTGTGGACAGCTGTACCTGCGAGTTATCTCCATCGTGGCAGCAGGACTGATCT 849
|||||
QY 809 TCGAGCTCTCTCTCTCTATTTCTCAAAATGCAATGAAGGGTTTGGGACTCCCGAGTGGTTCCTA 868
|||||
Db 850 TCGAGCTCTCTCTCTCTATTTCTCAAAATGCAATGAAGGGTTTGGGACTCCCGAGTGGTTCCTA 909
|||||
QY 869 TCTTGGCTGTAATCATATTTTCCATGACAAAATGAATGGGTAAAGGTGATTTCTATGATCC 928
|||||
Db 910 TCTTGGCTGTAATCATATTTTCCATGACAAAATGAATGGGTAAAGGTGATTTCTATGATCC 969
|||||
QY 929 TAGCTATTGGGGTTTCACTTCTCTATGCTACCAACAATATCTTGATGACAAAACCTTCA 988
|||||
Db 970 TAGCTATTGGGGTTTCACTTCTCTATGCTACCAACAATATCTTGATGACAAAACCTTCA 1029
|||||
QY 989 AGAAAAATCATGAAATCAACAACAGAAATCCCTGACCCACAGAGCAGAGAGTCAAA 1048
|||||
Db 1030 AGAAAAATCATGAAATCAACAACAGAAATCCCTGACCCACAGAGCAGAGAGTCAAA 1089
|||||
QY 1049 CTTGGCAATCAAAATAA 1065
|||||
Db 1090 CTTGGCAATCAAAATAA 1106
|||||

RESULT 6
AX033550
LOCUS Sequence 7 from Patent WO049152.
DEFINITION AX033550
ACCESSION AX033550
VERSION AX033550.1 GI:10280294
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Andre,B., Buerkle,L., Frommer,W.B. and Gillissen,B.
TITLE Nucleic acids that code for a nucleobase transporter
JOURNAL Patent: WO 0049152-A 7 24-AUG-2000;
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLFF B (DE) ;
GILLISSEN BERND (DE)
FEATURES Location/Qualifiers
1. .1071
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"

ORIGIN		/db_xref="taxon:3702"		Db		1040 GTGAGCAATCAAAATAAGCTGTTACTTCAAA 1070	
Query Match		77.3%; Score 835.8; DB 6; Length 1071;		RESULT 7			
Best Local Similarity		90.2%; Pred. No. 2.9e-194;		AX033547		1293 bp DNA linear PAT 21-SEP-2000	
Matches 894; Conservative 0; Mismatches 97; Indels 0; Gaps 0;				LOCUS			
				DEFINITION		Sequence 4 from Patent WO0049152.	
Qy		89 GAGTATCTATGACACACGAGGAAACAGTAATAATGGCTAGCAACGGTAGTTCAACTTGTG 148		AX033547			
Db		80 GACTATACTATGAAATGGAGGAAACAGCAAAATGGCTAGCAACGGTAGTTCACTTGTAG 139		VERSION		AX033547.1 GI:10280291	
				KEYWORDS			
Qy		149 GCTTTCCTGCTACTTCCATATATATCTTGTGTCATTTTAAACACATGCAACAACACTGATA 208		SOURCE		Arabidopsis thaliana (thale cress)	
Db		140 GCTTTCCTTACTTCCATATATCTTGTGTCATTTTAAACACATGCAACAACACTGAGA 199		ORGANISM		Arabidopsis thaliana	
				REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.	
Qy		209 GAGATGGAAGAACCTCACCTAGGAACCGGTATTTGGTTTACGTAGTGTGGACTTC 268		AUTHORS		Andre,B., Buerkle,L., Frommer,W.B. and Gillissen,B.	
Db		200 GAGATGGCAATTAACCTCACCTAGGAACCGGTATTTACATAGTGTGGACTTC 259		TITLE		Nucleic acids that code for a nucleobase transporter	
Qy		269 TTGTAGGACGAGTGTCTATCTGTACTCCATTTGACATTTCTTTACTTACCCGTTTCTAGCT 328		JOURNAL		Patent: WO 0049152-A 4 24-AUG-2000;	
Db		260 TTGTAGGACGAGTGTCTATCTATTTCCATTTGGACTGCTTTTACCTACCTGTTTCTACCC 319		FEATURES		ANDRE BRUNO (BE) ; BUEKLE LUKAS (DE) ; FROMMER WOLF B (DE) ; GILLISSEN BERND (DE)	
Qy		329 ATTCCCTCATCTGTGCATCTCAGTTAGCCTTCAATGCTTCTCTCTTATTTTCTTAAT 388		source		1. .1293	
Db		320 TTTCCCTGATCTGTGCATCAAGTTAGCCTTCAACGCTTCTCTCTTATTTTACTCAACT 379		/organism="Arabidopsis thaliana"			
				/mol_type="unassigned DNA"			
				/db_xref="taxon:3702"			
Qy		389 CACAAAACTTACCCCTATCTTTTAAATCTCTTTTCTCTTAACTATATCTTCCACCC 448		ORIGIN			
Db		380 CACAAAACTTACTCTATCATTTTGAATTTCTTTTCTCTCTCACTATATCTTCCACCC 439		Query Match		37.6%; Score 406.8; DB 6; Length 1293;	
Qy		449 TACTTGCATTCATTAATGAGGAGACAGACTCCACAAAAGTTTACAAAAGGAGATGTCA 508		Best Local Similarity		64.8%; Pred. No. 4.9e-89;	
Db		440 TCCTTGCATTTAATAACGAGGAATCAGATTCCAAAAAGTTTACAAAAGGAGATGTCA 499		Matches 622; Conservative 0; Mismatches 332; Indels 6; Gaps 1;			
Qy		509 AAGTTTTCATATGCACCGTTCGCTGCTGCTGTTATGTTGTTATGTTTATCCCTCAAC 568		Qy		89 GAGTATCTATGACACACGAGGAAACAGTAATAATGGCTAGCAACGGTAGTTCAACTTGTG 148	
Db		500 AAGTTTTCATATGCACCGTTCGCTGCTGCTGTTATGTTGTTATGTTTATCCCTCAAC 559		Db		293 GAGTTTACTATGAAAAATGGTGGGAAGAGTACATGGATGGGAACACTTGTCCAACTATCG 352	
Qy		569 AGCTAGCCTTTCTAAAGTCTCTAAAGAACGAAAAATTTCTCAGAAAGTTATGGATATGATA 628		Qy		149 GCTTTCCTGCTGCTACTTCCATATATATCTTGTTCATTTTAAACACATGCAACAACACTGATA 208	
Db		560 AGCTAGCCTTTCTGTAAGTTTAAAGAACGAAAACTTTCTCAGAAAGTTATATAATATGATA 619		Db		353 GCTTCCCTGTTGTTTCTCTTCGGTTCCTTTTCCAAACCAAAAATCCCAACCAACAG 412	
Qy		629 TCTAGCTGAGTCTAGTGGCGAGTTGTGTAGGTGGTGGGCTTTTGGTACGAGT 688		Qy		209 GAGATGGAAGAACCACTCACCT-----AGGAACCGTGTATTTGGTTTACGTAGTGTG 262	
Db		620 TCTACATGAGTCTAGTGGCCAGTTGTGTAGCGTGGTGGGCTTTTGGTCTAGTAGCGAGT 679		Db		413 AAGCAGATTTTCAGAAAGTCTCTTCCTTCACCATTTCTTGGATCAGTTTACATGTTACTG 472	
Qy		689 GGAAAACTTTGAGCAGTGAATGGGATAACTACAAACATGGGAAGGTATCCTACATTATGA 748		Qy		263 GACTTCTGTAGAGACAGATTGCTATCTGTACTCTCAATGAGACTTCTTACTTACCCGTTT 322	
Db		680 GGAAAACTTTGAGCAGTGAATGGGAAACTACAAACTTGGGAGGATTCCTATGTCTATGA 739		Db		473 GACTTATGAGTCTGTGCTAACTCTTATATGTCCTCTGTGTTGTACTACTACTACCACTTT 532	
Qy		749 ACCTAGTGTGGACAGCTGTTACCTGGCAGTTATTTCCATCGGTGGCACAGGACTGATCT 808		Qy		323 CTACCTATCCCTGATCTGTGCACTCTCAGTTAGCTTCAATGCTTCTCTCTTATTTTC 382	
Db		740 ACCTAGTGTGGACAGCTGTTACCTGGCAGGTTATTTCCATCGGTGGCACAGGACTGATCT 799		Db		533 CTACTTTCTCCTCACTTGGCCCTCACAATTTGGCCTTCACTGCTTCTCTCATATTTTC 592	
Qy		809 TCGAGCTCTCTCTATTTCAATATGCAATTAAGGTTTTGGGACTCCAGTGGTTCTCTA 868		Qy		383 TTAACTCACAACAACTTACCCCTATCATTTTTTAAATTTCTTTTCTCTCTTAACTATATCTT 442	
Db		800 TCGAGCTTTCTCTCCCTATTTCTCAATGCAATTAAGCGCTTTGGGACTCCCGTGGTTCTA 859		Db		593 TAAACTCGCAGAGTTTCACACCTTTTCATTGTGTAATTTCTGTCTCTCTTACTATTTCTCT 652	
Qy		869 TCTTGGCTGTATCATTTTCCATGACAAAAATGAATGGGTTTAAAGTGATTTCTATGATCC 928		Qy		443 CCACCTCTACTTGCATTCATTAATGAGGAGACAGACTCCACAAAAGTTTACAAAAGGAGT 502	
Db		860 TCTTGGCTGTATCATTTTCCATGACAAAATGAACGGCTTAAAGTGATTTCTATGATTC 919		Db		653 CTGCCCTCTCTGTTGTCACACCTGATTCGGAACACACAGCAAAAGTGTCTAGAGTAAAT 712	
Qy		929 TAGCTATTTGGGGTTTCACTTCTCTATGTCTACCAACAATATCTTGTGATGACAAAACTTGA 988		Qy		503 ATGTCAAAAGGTTTTCATATGACACCGTTGCTGCTGCTGCTGTTATGTTCTAGTCTTATCCC 562	
Db		920 TAGCTATTTGGGGTTTTCGTTATCTTATGTTCTACCAACAATATCTTGTGATGATTC 979		Db		713 ATGTGATGGGATNATATGATGACCATTTGCTGCTCTGCTGGGATTTGGATTTGCTGCTATCCC 772	
Qy		989 AGAAAAATCATGAATATCAACACAGAAATGCCCTGACCCACGACGACGACGAGTCAA 1048		Qy		563 TACAAACAGCTAGCCCTTTTAAAAAGTTCCTAAAGAGCAAAAAATTTCTCAGAAAGTTATGATA 622	
Db		980 AGAAAAAGTAATGAAATACCAACAGAAATCCCCTGACCCGACCGACGAGCAAGAGGGTCAA 1039		Db		773 TGGTACAACTGATCCTCAGGAAGGTTTTTAAAGAAGCAAAACATTTCTCAACGCTCACTGACT 832	
Qy		1049 CTGTGCAATCAAAATAAGCTGATATTTTGA 1079		Qy		623 TGATAATCTAGTGAGTCTAGTGGCCAGTTGTGTACCGTGGTGGGCTTTTCTGCTAGCA 682	
				Db		833 TGGTCGTTTACCAATCTCTAGTTTGCAGAGCTGTGTGGTCTCTCATAGGAGCTTTTTCGAAGCG 892	
				Qy		683 GTGAGTGGAAAACTTTTGAGCAGTGAATGGATAACTACAAAACATGGGAAGGTATCTCTACA 742	

```
Db      893  GGGAGTGGAAACTTTAAACAAGTGAGATGGAACCTACAAACTGGGGAAGTGCCATACG 952
QY      743  TTATGACCTAGTGTGGACAGCTGTTACCTGGCAGTTATTCTCCATCGGTGGCAGGAC 802
Db      953  TTATGACTTTGGCTCGTAGCTATTCTCTGGCAAGCTTACACCATTTGGCTGTGGGAC 1012
QY      803  TGATCTTCGAGCTCTCTCTATTCCTCAATCAATAGCGTTTGGGACTCCCAAGTGG 862
Db      1013  TGATCTTTGAGTCATCTCTGTGTTCTCCAATTCCTCAATCTGTGTGGGATTGCCTATAG 1072
QY      863  TTCTATCTTGGCTGTAATCATTTTCCATGACAAATGAATGGTTAAAGGTCATTCTA 922
Db      1073  TTCCAGTTGTAGCAGTGATGTTTTCATGATAAAATGAACCGCTCAAAAGATCTTCTCCA 1132
QY      923  TGATCTAGCTATTTTGGGGTTTCACATCTCTATGTTACCAACAATATCTTTGATGACAAA 982
Db      1133  TCATTTTAGCTATCTGGGATTCATTTCATTTGCTATCAGACTACCTCGAGGAAAGA 1192
QY      983  ACTTGAAGAAAATCATGAAATCACAAACAGAAATCCCTGACCCACAGAGAGAG 1042
Db      1193  AGTTGAAGACTAGCCACACAAGTCTCTGTAGGAGATCTCTATCTACTACCTGTGAGGAAG 1252

RESULT 8
LOCUS   ATF15J5
DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone F15J5 (ESSA project).
ACCESSION AL110123
VERSION   AL110123.1
KEYWORDS  GI:5816989
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1
Bevan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H.W.,
Mayer, K.F.X., Lemcke, K. and Schueller, C.
Unpublished
2 (bases 1 to 58427)
EU Arabidopsis sequencing project.
Direct Submission
Submitted (30-MAR-1999) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG. E-mail:
schueller@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bbsrc.ac.uk
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.

FEATURES             Location/Qualifiers
     source           1..58427
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosomes="4"
     misc_feature     1..2548
                     /note="overlap to BAC T6K21; please refer to EMBL:AL021889
                     for analysis and annotation"
     gene            3596..5023
                     /gene="F15J5.10"
                     complement(join(3596..3646,3736..3801,3895..4020,
                     4116..4281,4725..5023))
     gene            4116..4281
                     /gene="F15J5.10"
                     complement(join(3596..3646,3736..3801,3895..4020,
                     4116..4281,4725..5023))
     CDS             4116..4281
                     /gene="F15J5.10"
                     /note="Contains Eukaryotic initiation factor 4E signature
                     (IDFKRHIIEPKWEDPICANGKW)"

contains EST gb:T21435, AA586111"
/codon_start=1
/product="translation initiation factor eIF4E"
/protein_id="CAB51645.1"
/db_xref="GI:5816990"
/db_xref="GOA:O23252"
/db_xref="Swiss-Prot:O23252"
/translation="MAVEDTPKSVVTEAKPNSIENPIDRVHEEGDDAEBEGIAGGEG
DGNVDESSKGVPSHPLEHSGTWTFWDPNPAVKSKOTSGSSLRPVTFTSTVEFWSLY
NMKHPSKLAGADPYCPKHLIEPKWEDPICANGGKWTMTTPKKSCKSLVTLALAI
GEQPDGDEICGAVNIRKQERISITWTKNASNEAAQVSIQKWKEFLDYNNSIGFII
HEDAKLDRNAKNAYTA"
/complement(3596..3646)
/gene="F15J5.10"
/number=1
complement(3647..3735)
/gene="F15J5.10"
/number=1
complement(3736..3801)
/gene="F15J5.10"
/number=2
complement(3802..3894)
/gene="F15J5.10"
/number=2
complement(3895..4020)
/gene="F15J5.10"
/number=3
complement(4021..4115)
/gene="F15J5.10"
/number=3
complement(4116..4281)
/gene="F15J5.10"
/number=4
complement(4282..4724)
/gene="F15J5.10"
/number=4
complement(4725..5023)
/gene="F15J5.10"
/number=5
6767..6789
9077..13984
/gene="F15J5.20"
join(9077..9301,9805..9859,9938..10113,10198..10419,
10487..10725,10806..11331,11419..11692,11793..12736,
12827..13093,13169..13432,13508..13984)
/gene="F15J5.20"
/note="similarity to multidrug resistance protein 1,
Aspergillus fumigatus, U62933
Contains ABC transporters family signature,
ABC Transporter [MSGGQKQRIAIARAI][LUSGQKQRIAIARAI],
ATP/GTP-binding site motif A (P-loop), Atg_Gcp_A
[QSGSGSKS][GESGSGKS]"
/codon_start=1
/product="multidrug resistance
protein/P-glycoprotein-like"
/protein_id="CAB51646.1"
/db_xref="GI:5816991"
/db_xref="GOA:Q9SVW6"
/db_xref="TrEMBL:Q9SVW6"
/translation="WEEKSSKKNDGGNQKVSFFKLFSFADKTDVVLMTVGTIAAAGNG
LTQPFMTLIFQQLINAFGTDPDHPMREVKVAVKFIYLAIVSCVAFQVSCMMVTG
ERQATIRGLYKLTQIDIGYFDTETNGVIGRMSGDTLLIODMGEKVKFTQLL
CTFLGGFAIAPYKGPLLAGLVCSCPLIVIAGAASLMSKMGAGQVAYAEAGNVE
QTGAIITVVAFTGEKQNTKEYSKLEIAYTVVQGLISGFGLTMLAVIFCSGLA
WYGAKLIMEKGYGGQVIVFAVLTTGMSLQDTSQSLNAPFAAGRAAFKMFETIKR
SPKIDAYDMSGVLELDIQLDQVYFPAFDVQIFAGSLFVPGKTVALLVQGS
GSGKSTVSLIERFYDESGQVLIDNIDKLQKIRSKISGLVSGEPLPATTIKEN
IAYGEDATDOEIRTAELANAAPFIDKLPOGLTMVGEHGTOMSGQKQALAIARAI
LKNPKILLDEATSDAESRIVQDALNLMNSRTTVVAHRLTTITFADVIIVVHO
KRIVKGTHEMIQDPGQNSQLVRLQGSKEATESETPETSLDVERSGSLRSSKPE
RVSVRSNRRSHRSLSANMFFPGVNGQVDEENNVRHKKSLHFWALYIALGJTNF
IPVLVLSIAAMVHGTVPIFGLLSSINNFYPAKILKKDLSHFWALYIALGJTNF
VMIPVPVYFFGIAGGKLIKIRSMCFDKVHQEISWFDFTANSRYIYIINRRLY
```



```
Db 57762 AGTTGAAGACTAGCCACACAGTCTCTGTAGGAGATCCTCTACTACCTGCTGTGAGAAG 57821
|||||
RESULT 9
AX507517 3387 bp DNA linear PAT 27-SEP-2002
LOCUS
DEFINITION
Sequence 2212 from Patent WO0216655.
ACCESSION
AX507517
VERSION
AX507517.1 GI:23388754
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
1
AUTHORS
Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
TITLE
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL
Patent: WO 0216655-A 2212 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)
FEATURES
Location/Qualifiers
source
1..3387
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 37.6%; Score 406.2; DB 6; Length 3387;
Best Local Similarity 64.7%; Pred. No. 6.5e-89;
Matches 622; Conservative 0; Mismatches 333; Indels 6; Gaps 1;
QY 88 GGAGTATCTATGACAAACGAGTAAATGGCTAGCAACGGTAGTTCAACTGTT 147
|||||
Db 226 GTAGTTTACTATGAAATGGTGGGAAGATGATGATGGAACACTTGTCACCTAATC 285
|||||
QY 148 GGCTTTCTGTGCTACTTCCATATATATCTTCTCATTTTAAACACATGCAACACTGAT 207
|||||
Db 286 GGCTTCCCTGTTCTGTTCTTCTCCGCTTCTTTCCCAACCAAAATCCCAACCAACA 345
|||||
QY 208 AGAGATGGAAAGAAAGAACTCACCT-----AGAACCGTGTATGTTTACGTAGTGGTT 261
|||||
Db 346 GAAGCAGATTTCAGAAAGTTCTCTCTCCACCATTTCTGGATCAGTTTACATCGTTACT 405
|||||
QY 262 GGACTTCTGTAGGACGAGATGCTATCTGCTACTCCATTGGAGTCTTTTACTTACCGTT 321
|||||
Db 406 GGACTATAGTGTCTGCTAACTCTTATATGTCCTCTGTTGGTTTACTATACTTACCAGTT 465
|||||
QY 322 TCTACCTATTCCCTGATCTGTGCATCTCAGTTAGCCCTTCAATGCTTTCTTCTTATTTT 381
|||||
Db 466 TCTACTTTCTCCCTCATCTTGGCTCACAATGGCCCTTCACTGCCCTTCTTCTCATATTTT 525
|||||
QY 382 CTTAACCTCAGAAACCTTACCCCTATCATTTTAAATCTCTTTTCTCTTAACTATATCT 441
|||||
Db 526 CTAACCTCGCAGAAATTCACACCTTTCAATGTGAATTTCTGTCTTCTCTTACTATTTCC 585
|||||
QY 442 TCCACCCCTCTTCATTCATATATAGGACAGACTCCACAAAAGTTACAAAAGGAGAG 501
|||||
Db 586 CTGCCCCCTCTCTGGTGTCAACTGTATTCGAAAACACAGAAAAGTGTCTAGAGTAAA 645
|||||
QY 502 TATGTCAAAAGTTTCATATGACCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
|||||
Db 646 TATGTGATAGGATATATGTACCATTTGGTGTCTTCTGCTGGGATTTGGATTTGCTATCC 705
|||||
QY 562 CTACACAGCTAGCTTTCTTAAAGTCTTAAAGACGAAATTTCTCAGAAAGTTATGGAT 621
|||||
Db 706 CTGGTCAACTGATCTCTCAGGAAGGTTTTAAAGAACGAAACATTTCTCAACGGTCACTGAC 765
|||||
QY 622 ATGATAATCTAGTGTAGTGTAGTGGCCAGTTGTTAGCGTGGTGGGCTTTTGTGAC 681
|||||
Db 766 TTGGTCGCTTACCAATCTCTAGTTGCAAGCTGTGGTTCTCTATAGGACTTTTTCGCAAGC 825
|||||
```

```
QY 682 AGTGRATGGAAACCTTTGAGCAGTGAATGATTAACATAACATGGGAGGTATCTCTAC 741
|||||
Db 826 GGGAGTGGAAACCTTTTAAACAGTGAATGGAATACTACAACTGGGGAAGTGCCTATC 885
|||||
QY 742 ATTATGAACCTAGTGTGGACGCTGTCTACCTGGCAGTATTCTCCATCGGTGGCAGGA 801
|||||
Db 886 GTTATGACTTTTGGCCCTCGATAGCTATTCTTCGGCAAGCTACACCAATGGCGCTGCGG 945
|||||
QY 802 CTGATCTTCGAGCTCTCTCTCTTATTCTCAAAATGCAATAAGCGTTTTTGGACTCCCGATG 861
|||||
Db 946 CTGAATCTTGTAGTCATCTTCTGTCTCTCAATTCATTAATCTGCTGGGATTCCTATA 1005
|||||
QY 862 GTTCTTATCTTGGCTGTAAATCAATTTTCCATGACAAATGAATGGGTAAAGGTGATTTCT 921
|||||
Db 1006 GTTCCAGTTGTAGCAGTGATTGTTTTCCATGATATAAATGAACGCGTCAAAAGATCTTCTCC 1065
|||||
QY 922 ATGATCTAGCTATTGTGGGTTTCACTTCTATGTCTACCAACAATATCTTATGACAAA 981
|||||
Db 1066 ATCATTTTGTAGCTATCTGGGGATTTCATTTCTATCTATCAGCACTACCTCGACGAAAAG 1125
|||||
QY 982 AACTTGAAGAAAATCATGAAATCAACAACAGAAATCCCTGACCCACCAAGACAGAA 1041
|||||
Db 1126 AAGTTGAAGACTAGCCACAAAGTCTGTAGGAGATCTCTACTACTACTGCTGAGGAA 1185
|||||
QY 1042 G 1042
Db 1186 G 1186
|||||
RESULT 10
AX652015 3387 bp DNA linear PAT 22-MAR-2003
LOCUS
DEFINITION
Sequence 908 from Patent WO03000898.
ACCESSION
AX652015
VERSION
AX652015.1 GI:29154833
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
1
AUTHORS
Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katagiri, F., Quan, S., Tao, Y., Whitcham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE
Plant genes involved in defense against pathogens
JOURNAL
Patent: WO 03000898-A 908 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
Location/Qualifiers
source
1..3387
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 37.6%; Score 406.2; DB 6; Length 3387;
Best Local Similarity 64.7%; Pred. No. 6.5e-89;
Matches 622; Conservative 0; Mismatches 333; Indels 6; Gaps 1;
QY 88 GGAGTATCTATGACAAACGAGGAAACAGTAAATGGCTAGCAACGGTAGTTCAACTGTT 147
|||||
Db 226 GTAGTTTACTATGAAATGGTGGGAAGATGATGATGGAACACTTGTCACCTAATC 285
|||||
QY 148 GGCTTTCTGTGCTACTTCCATATATATCTTGTCTATTAAACACATGCAACACTGAT 207
|||||
Db 286 GGCTTCCCTGTTCTGTTCTTCTTCCGCTTCTTTCCCAACCAAAATCCCAACCAACA 345
|||||
QY 208 AGAGATGGAAAGAAAGAACTCACCT-----AGAACCGTGTATGTTTACGTAGTGGTT 261
|||||
Db 346 GAAGCAGATTTCAGAAAGTTCTCTCTCCACCATTTCTGGATCAGTTTACATCGTTACT 405
|||||
QY 262 GGACTTCTGTAGGACGAGATGCTATCTGCTACTCCATTGGAGTCTTTTACTTACCGTT 321
|||||
Db 406 GGACTATAGTGTCTGCTAACTCTTATATGTCCTCTGTTGGTTTACTATACTTACCAGTT 465
|||||
QY 322 TCTACCTATTCCCTGATCTGTGCATCTCAGTTAGCCCTTCAATGCTTTCTTCTTATTTT 381
|||||
Db 466 TCTACTTTCTCCCTCATCTTGGCTCACAATGGCCCTTCACTGCCCTTCTTCTCATATTTT 525
|||||
QY 382 CTTAACCTCAGAAACCTTACCCCTATCATTTTAAATCTCTTTTCTCTTAACTATATCT 441
|||||
Db 526 CTAACCTCGCAGAAATTCACACCTTTCAATGTGAATTTCTGTCTTCTCTTACTATTTCC 585
|||||
QY 442 TCCACCCCTCTTCATTCATATATAGGACAGACTCCACAAAAGTTACAAAAGGAGAG 501
|||||
Db 586 CTGCCCCCTCTCTGGTGTCAACTGTATTCGAAAACACAGAAAAGTGTCTAGAGTAAA 645
|||||
QY 502 TATGTCAAAAGTTTCATATGACCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
|||||
Db 646 TATGTGATAGGATATATGTACCATTTGGTGTCTTCTGCTGGGATTTGGATTTGCTATCC 705
|||||
QY 562 CTACACAGCTAGCTTTCTTAAAGTCTTAAAGACGAAATTTCTCAGAAAGTTATGGAT 621
|||||
Db 706 CTGGTCAACTGATCTCTCAGGAAGGTTTTAAAGAACGAAACATTTCTCAACGGTCACTGAC 765
|||||
QY 622 ATGATAATCTAGTGTAGTGTAGTGGCCAGTTGTTAGCGTGGTGGGCTTTTGTGAC 681
|||||
Db 766 TTGGTCGCTTACCAATCTCTAGTTGCAAGCTGTGGTTCTCTATAGGACTTTTTCGCAAGC 825
|||||
```


Db 406 GCATATTAGTGTCTGAACCTTTATATATGTCTCTGTGGTTTACTATATCTTACCAGTT 465
Qy 322 TCTACTATTCCTGATCTGTGCATCTCAGTTAGCTTCAATGCTTTCTTCTTATTTT 381
Db 466 TCTACTTTCTCCTCATCTTGGCCCTCAAAATTTGGCCTTCACTGCTTCTTCTCATATTTT 525
Qy 382 CTTAACTCACAAAACTTACCCTATCATTTTAAATTTCTTTTCTTCTTAACTATATCT 441
Db 526 CTAACTCCGAGAAGTTACACCTTTCAATGTGAATTTCTGTCTTCTTACTATTTCC 585
Qy 442 TCCACCTACTTGCAATCAATATGAGGAGACAGACTCCAAAAAGTTTACAAAAGGAGAG 501
Db 586 TCTGCCCTCCTCGTGTCAACACTGATTCGGAACACACAGAAAAGTGTCTAGAGTAAA 645
Qy 502 TATGTCNAAAGTTTCATATGACCCCTTGTGCTGTCTGTTATGTTAGTCTTAGTCTTATCC 561
Db 646 TATGTGATAGGATAATATGTAACCATTTGTTCTTCTGCTGGATTTGGATTTGCTGCTATCC 705
Qy 562 CTACAAACAGCTAGCCTTTCTAAAGTCTTAAAGAGCAAAATTTCTCAGAAGTTATGGAT 621
Db 706 CTGTTACAACTGATCCTCAGNAGGTTTAAAGAGCAAACTTCTCAACGGTCACTGAC 765
Qy 622 ATGATAATCTACGTGAGTCTAGTGCCAGTTGTGTAGCGTGGGCTTTTGTCTAGC 681
Db 766 TTGTCGCTTACCAATCTCTAGTTGCAAGCTGTGTGTTCTCATAGGACTTTTCGCAAGC 825
Qy 682 AGTGAGTGGMAAATTTGAGCAGTGAATGATGATCACTCAAACTGGAAGGTATCCTAC 741
Db 826 GGGGAGTGMAAATTTTAAAGTGAGATGGAATACTCAAACTGGGGAAGTGCCTAC 885
Qy 742 ATTATGAACCTAGTGTGACAGCTGTATCTGCGAGTTATTTCTCATCGTGGCAGGA 801
Db 886 GTATGACTTTGGCCTCGATAGTATTTCTGCGCAAGTCTACACATTTGGGCTGTGGA 945
Qy 802 CTGATCTTCGAGCTCTCTCTCTATCTCAAAATGCAATGAGCGTTTGGGACTCCAGTG 861
Db 946 CTGATCTTTGAGTCATCTCTGTGTCTTCCAAATTCATAAATGCTGTGGGATTGCTATA 1005
Qy 862 GTTCTATCTTGGCTGATCATCTTTTCCATGACAAATGAATGGTTAAAGTGATTTCT 921
Db 1066 GTTCCAGTTGTAGCAGTGAATTTTCCATGATAAAATGAACGGGTCAAAAGTCTTCTCC 1065
Qy 922 ATGATCTAGCTATTTGGGGTTTCACTTCTATGTCTACCAACATATCTTTGATGACAAA 981
Db 1066 ATCATTTAGCTATCTGGGATTCATTTCTATCTATCAGCACTACTCGACGAAAAG 1125
Qy 982 AACTTGAAGAAAATCATGAATCAACACAGAAATCCCTGACCCACCAAGACAGAA 1041
Db 1126 AAGTTGAAGACTAGCCACACAAGTCTGTAGGAGATCCTCATCTACTACTGCTGAGGAA 1185
Qy 1042 G 1042
Db 1186 G 1186

RESULT 11
AP006383/c
LOCUS
DEFINITION
Lotus corniculatus var. japonicus genomic DNA, chromosome 6,
clone:ljT05120, TM0245, complete sequence.
ACCESSION
AP006383
VERSION
AP006383.1 GI:31581014
KEYWORDS
HTG
SOURCE
Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM
Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
1
REFERENCE
Asamizu,E., Kato,T., Sato,S., Nakamura,Y., Kaneko,T. and Tabata,S.
AUTHORS
Structural Analysis of a Lotus japonicus Genome. IV. Sequence
TITLE
Features and Mapping of seventy-three TAC clones which cover the

7.5 Mb Regions of the Genome
DNA Res. (2003) In press
2 (bases 1 to 73179)
Sato,S.
Direct Submission
Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-Kamatari, Kisarazu,
Chiba 292-0818, Japan (E-mail:ssat@kazusa.or.jp,
URL: http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),
Fax:81-438-52-3934)
Location/Qualifiers
1. 73179
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="6"
/clone="LjT05120"
/clone_lib="LjT library"
/note="TAC clone:TM0245-synonym: Lotus japonicus"
ORIGIN
Query Match 35.6%; Score 385.2; DB 8; Length 73179;
Best Local Similarity 63.7%; Pred. No. 7.7e-84;
Matches 604; Conservative 0; Mismatches 338; Indels 6; Gaps 1;
Qy 58 CACATACAAACGGTGGCTCAGGAGTATCTATGCAACGGAGGAAACAGTAAATGGCTAG 127
Db 64772 CAGCTACAAACCCCTTTTGGGAAGATTGTACTATGAAATGGTGAAGAGCAATGGATGG 64713
Qy 128 CAACGGTAGTTCAACTTGTGGCTTTCTGTGCTACTTCCATATATATATCTTTGTCATTTA 187
Db 64712 CAACACTGTGCCAGTTGTGGTTTCCAAATCTGTGCTCTATATACATCTCAGCAC 64653
Qy 188 AAACACATGCAACAACTGATAGAGATGGAAGAAAGAACCTCACCTPAGAAACCGTGT- 242
Db 64652 CCAAAGATCTCACTGAAAACAGCTTTTATCATCAGAATCAGCCTTCTAAATTAGTGCTAA 64593
Qy 243 -ATTGGTTAGGTAGTGTGGACTTCTTGTAGGAGCAGATTGCTATCTGTACTCCATTG 301
Db 64592 CACTTGTCTATGTCTCCCTTGGCCAACTTGTGACATTTAAATTTGTTATTTAAATTTCTGCTG 64533
Qy 302 GACTTCTTTACTTACCCTTTTCTACCTATTCCCTGTGATCTGTGCATCTCAGTTAGCCTTCA 361
Db 64532 GGCTCTGGTTCCTTCCAGTCTCTACTATTTCATCTATTGTTTCACTCAGTTGGCTTTCA 64473
Qy 362 ATGCTTTTCTCTCTTATTTTCTTAACTCAGAAAACCTTACCCCTATCATTTTAAATTTCTC 421
Db 64472 ATGCTATTTTCTCCTACTTCTCTCAATTCACCTCAAGATCAGACCTTATCATCATCACTCTC 64413
Qy 422 TTTTCTCTTAACTATATCTTCCACCTACTTGCATTTCAATATGAGGAGCAGACTCCA 481
Db 64412 TAGTCTTCTTAAACCAATTTCTCTGCTTCTGCTTGTGTTCAAGATGAGTCACTGATTTCAA 64353
Qy 482 CAAAAGTTTACAAAAGGAGAGTGTCTCAAGGTTTTCATATGACCGTGTCTGCTGCTG 541
Db 64352 CACAAATCTCCAAACAAAAGTATGTCAATTGGATTCATATTCACAGTTGCTGCATCAGCTG 64293
Qy 542 GTTATGCTGTAGTCTTATCCCTACACAGCTAGCCTTTCTAAAAAGTCTCTAAAGAGCAAA 601
Db 64292 GGGTGGATTAGGACTTTTCCCTTTACACAGTTTGTCTTCAAGAACGTCATAAAAAGGAAA 64233
Qy 602 ATTCTCAGAAGTTATGGATATGATATCTACGTGAGTCTAGTGGCCAGTTGTGTTAGCG 661
Db 64232 ATTTCAAAGTGCATAGATATGATAGTTTATCTGTCCCTTTGTAGCTACATGTGCTAGTC 64173
Qy 662 TGGTGGGCTTTTGTCTAGCAGTGTAGTGAAGAACTTTGAGCAGTGAATGGATAACTACA 721
Db 64172 TAGTGGGACTTTTGTCTAGTGGAGAGTGAAGAGTTTGAAGAAAGAAATTTGAGGGGTATG 64113
Qy 722 AACATGGGAAGGTATCCTCATATATGAACCTAGTGTGGACAGCTGTTACCTGGCAGTTAT 781
Db 64112 AGACAGGTAAAGCTGCCTATATATATGAAATCTTGTCTTTTACAGCCATATTTTCGCAACTCG 64053

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-Kamatari, Kisarazu,
Chiba 292-0818, Japan (E-mail:ssat@kazusa.or.jp,
URL: http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),
Fax:81-438-52-3934)
Location/Qualifiers
1. 73179
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="6"
/clone="LjT05120"
/clone_lib="LjT library"
/note="TAC clone:TM0245-synonym: Lotus japonicus"
ORIGIN


```
QY 782 TCTCCATCGGTGGCAGAGCACTGATCTTCGAGCTCTCCTCTCTATTTCTCAATGCAATAA 841
Db 64052 TTACTCTCGGTGTGACGGCACTGGTTTTCGAGGCTCTTCCCTCTCTTCTCCAAATGCTATAA 63993
QY 842 GCGTTTTGGGACCTCCAGTGGTTCCTATCTTGGCTGTATCATTTTCCATGACAAATGA 901
Db 63992 GCGTTTTGGGTGACGATATTTGTCCTATATATGCTGTAAATGGTCTTACAGACAGAAATGC 63933
QY 902 ATGGGTTAAAGTGATTTCTTATGATCCTAGCTATTTGGGGTTTCACTTCTATGTCTACC 961
Db 63932 ATGGGATAAAGTCATTTCTATGCTGTGGCCGTTTGGGGTTTATATCGTACGTGTATC 63873
QY 962 ACAAATATCTTGATGACAAAACTTGAAGAAAATCATGAATACAA 1009
Db 63872 AGCAATACCTGGATGAATGCAATCCAAACAGAGATAGAGACAGAA 63825

RESULT 12
AX033548
LOCUS AX033548 1194 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 5 from Patent WO0049152.
ACCESSION AX033548
VERSION AX033548.1 GI:10280292
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
1
REFERENCE
AUTHORS Andre,B., Buerkle,L., Frommer,W.B. and Gillissen,B.
TITLE Nucleic acids that code for a nucleobase transporter
JOURNAL Patent: WO 0049152-A 5 24-AUG-2000;
ANDRE BRUNO (BE) ; BUERKLE LUKAS (DE) ; FROMMER WOLFF B (DE) ;
GILLISSEN BERND (DE)
FEATURES
source Location/Qualifiers
1..1194
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 34.2%; Score 369.6; DB 6; Length 1194;
Best Local Similarity 62.2%; Pred. No. 6.6e-80;
Matches 601; Conservative 0; Mismatches 359; Indels 6; Gaps 1;
QY 89 GAGTATCTATGACACGAGGAGAAACAGTAATGGCTAGCAACGGTAGTTCAACTTTGTG 148
Db 197 GATTGTACTATGAAATGAGGAGAAATAGCACATATGTGGTAACACTTCTTCAACTCATTG 256
QY 149 GCTTTCCTGCTGCTTCCATATATATCTTGTCATTTA-----AACACATGCAACAA 202
Db 257 GCTTCCCTGCTGCTGCTTCCGCTTCTTTCTCGATCAGGCAACCAATCAACAG 316
QY 203 CTGATAGAGATGAAAGAAACCTCACCTAGGAACCGTGTATTGGTTTACGTAGTCTTG 262
Db 317 ATACAAATTTCACTGCTGCTTCCCTTCTTCCACCCTTGCATCGGTTTACTTGTGCACTG 376
QY 263 GACTTCTTTAGAGAGAGATTTGCTATCTGTACTTCCATCGGACTTCTTTACTTACCGGTTT 322
Db 377 GACTGCTAGTGTCCGCTTATGCTTATTTGCTGCACTAGGTTGCTTACTTACCAGTCT 436
QY 323 CTACCTTATTCCTGATCTGTCGATCTCAGTTAGCCTTCAATGCTTCTCTCTCTATTTTC 382
Db 437 CTACTTCTCCCTCATCTTGGCTTCACAGTGGCTTTCAGTGCCTTTTCTCATATTTCC 496
QY 383 TTAACCTCAAAAACCTTACCCCTATCATTTTAAATCTCTTTTCTCTTAACTATATCTT 442
Db 497 TTAACCTCGAAAAGTTCACTCCCTTTGATAGTCACTTCTTGTCTTCTCTCACTGTATCT 556
QY 443 CCACCCCTACTTGCATTTCAATATAGGAGACAGACTCCCAAAAGTTACAAAAGGAGT 502
```

```
Db 557 CTGCTCTCTTGTGFTCAACACTGATTCAGAAAACCTCAATATGATATCTAGAGTACAGT 616
QY 503 ATGTCAAAGGTTTTCATATGACACCGTGTGCTGCTGCTGTTATGTTCTAGTCTTATCCC 562
Db 617 ATGTGATCGGTTTCATATGATACCATCGGTGCTTCCGCTGGGATTTGACTTGTACTATCTC 676
QY 563 TACAACAGCTAGCCTTTCTTAAAGTCTTAAAGACCAAAATTTCTCAGAAGTTATGGATA 622
Db 677 TGATACAAATGCTCTTTCAGGAAAGTTTTCACGAAGCATACATCTCAGCAGTCAACGACT 736
QY 623 TGATATATCTACGTAGTCTAGTGGCCAGTGTGTTAGCGTGGTGGGCTTTTGTCTAGCA 682
Db 737 TGCCCATTTTACCAGTCTCTAGTTGGAGTTGTGTAGTTCTCATAGGACTTTTTTGCAGTG 796
QY 683 GTGAGTGGAAAACTTTTGAGCAGTGAATGGATAACTACAAACATGGGAAGGTATCTTACA 742
Db 797 GAGAGTGGGAACTTTTGGCAAGTGAGATGAGAAACTACAAACTCGGAAAGTGTCTATATG 856
QY 743 TTATGAACCTAGTGTGGACAGCTGTTACCTGCGAGCTTATTTCTCCATCGGTGGCAGGAC 802
Db 857 TTTTGACTTTTAGCCTCGGCAGCTATTTCTCTGGCAAGTCTACACTCTTGTCTTGTGGAT 916
QY 803 TGATCTTTCGAGCTCTCCTCTCTATTTCTCAAATGCAATAGCGTTTGGGACTCCCGAGTGG 862
Db 917 TGATCTTTCGAGTCACTCCTCTGTGTTCTCCAATTCATTAACAGCTGTGGGATTTGCCTATAG 976
QY 863 TTCTCTATCTTGGCTGTAATCAATTTTCCATGACAAAATGAATGGGTTAAAGGTTATTTCTA 922
Db 977 TTCCAGTTTGGCAGCTAGTATTTTCCATGATAGATGAGCGCATCAAAAATTTCTTCCA 1036
QY 923 TGATCCTAGCTATTTGGGTTTCACTTCTCTATGTTTACCAACAATATCTTGTATGACAAA 982
Db 1037 TTTATTTTGTAGTATCTCGGCTTCTTCTTCTATGCTCTACAGCACTACCTCGACGAAAGA 1096
QY 983 ACTTGAAGAAAATCATGAATCAACAACAGATCCCTGACCCACCAAGAGAGAG 1042
Db 1097 AGTTGAATACCTAGCCACACAAAGTGTGTAGGAGATCTTCTATCTCTGTTAGGAAGGTC 1156
QY 1043 AGTCAA 1048
Db 1157 ACACAA 1162
```

RESULT 13

```
AY096558
LOCUS AY096558 1165 bp mRNA linear PLN 18-SEP-2002
DEFINITION Arabidopsis thaliana putative protein (At4g18200) mRNA, complete cds.
ACCESSION AY096558
VERSION AY096558.1 GI:20465496
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
1 (bases 1 to 1165)
Yanada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Open Reading Frame (ORF) Clones
2 (bases 1 to 1165)
Yanada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
```

Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE JOURNAL

Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

COMMENT

The Salk, Stanford, PGE (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Bath, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGE) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGE) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.

FEATURES source

Location/Qualifiers
1..1165
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U18628"
/ecotype="Columbia"
/note="This clone is in pUNI 51."
1..1165
/gene="At4g18200"
1..1134
/gene="At4g18200"
/codon_start=1
/evidence=experimental
/product="putative protein"
/protein_id="RAM20208.1"
/db_xref="GI:20465497"
/translation="MGISQVHYCNGDQNLNLDHETESFSPQTKNKRWLRSVI
YAIFFQPLATVGRLYYENGKSTYVTLQILGPPVILIFRFSRIRQPKSTDT
NFSQSPSTTLAVVLCGLVSAVLSAVGLLYLPVSTFSLILASQIAPTFYF
LNSQKFTPLIVNSLFLVSSALLVNTDSENTNVRQVIGFICTIGASAGIGLV
LSLQLLFRKVFTRHTSSAVLDLANYQSLVATCVVLGLFASGWRILPSEMRNKLK
KVSYLTLASAAIFMQVYTVGCVGLIFESSVFNSTITAVGLPIPVVAVIVFHKMD
ASKFSIILAIWGLSFVYQHYLDEKLLKTCQRPVEETQL"
1135..1165
/gene="At4g18200"

3' UTR

ORIGIN

Query Match 34.1%; Score 368.2; DB 8; Length 1165;
Best Local Similarity 62.8%; Pred. No. 1.5e-79;
Matches 590; Conservative 0; Mismatches 343; Indels 6; Gaps 1;
QY 59 TCACCTCTACACATCAACGGTGGCTCAGGAGTAGTACTATGACAAACGGAGAACAGTA 118
DB 152 TTGCGAACCACTTGGTACAGTTCTGGGTAGACTGTACTATGAAATGGAGGAAAGCA 211
QY 119 AATGCTAGCAACGGTAGTTCAACTTGTTGGCTTTCCTGTGCTACTCCATATTATATCT 178
DB 212 CATATGTGTTACACTTCTTCACTCACTTGCTTCCCTGCTACTGATTCTCTCGCTTCT 271
QY 179 TGTCAATTTA-----AACAATGCAACACTGATAGAGATGGAAAAAGACCTCACCTA 232
DB 272 TTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCCTTCTTCA 331
QY 233 GGAACCGTGTATGTTTACGTAGTGTGGACTTCTGTAGGACGACGATTGCTATCTGT 292

Db 332 CCACCCCTTGGCATCGTTTACTTGTGCACTGAGCTGCTAGTGTCTGCTTATTTGT 391
QY 293 ACTCCATTTGACCTTTTACTTACCGGTTTCTACCTATTCCTGATCTGTGCACTCAGT 352
Db 392 CTGCAGTTGGGTGCTTACTTACCACTCTACTTCTCCCTCATCTGGCCCTCACAGT 451
QY 353 TAGCCTTCAATGCTTCTTCTTATTTCTTAACTACACAAAACCTTACCCCTATCATTT 412
Db 452 TGGCCTTCACTGCGCTTTTCTCATATTTCTTAACTCGCAAAAGTTTCACCTTTGATAG 511
QY 413 TAAATCTCTTTTCTTAACTATATCTTCCACCTTCTTAACTCAATAAATGAGGAGA 472
Db 512 TCAATCTTGTGTTCTTCTTACGGTTTCTTCCCTCTGCGCTCTCTGCTCAACACTGATTCAG 571
QY 473 CAGACTCCACAAAGTTTACAAAAGGAGATGATGTAAGGTTTCAATGACCGTGTGCTG 532
Db 572 AANAACACAAATGTTATCTAGAGTACAAATGATGATGTTGGGTTTCATCTGTACCACTTGGTG 631
QY 533 CGTCTGCTGGTTATGGTCTAGTCTTATCCCTACACAGCTAGCTTCTTAAAGTCTCTAA 592
Db 632 CTTCCGCTGGGATTTGGATTTGGTACTATCTCTGATACAACTGCTCTCAGGAAAGTTTCA 691
QY 593 AGAAGCAAAATTTCTCAGAAAGTTATGATATGATTAATCTACGTGAGTCTAGTGGCCAGTT 652
Db 692 CGAAGCATACATCTCTAGCAGTCTCTGGACTTGGCAATTAACAGTCTCTAGTTGCAACTT 751
QY 653 GTGTTAGCGTGGTGGGCTTTTCTAGCAGTGTAGTGGGAAACCTTTGAGCAGTGAATGG 712
Db 752 GTGTGGTACTCATAGGACTGTTTGCAGAGTGGAGAGTGGAGAACTCTGCCAAGTGGATGA 811
QY 713 ATAACTCAACATGCGAAGTATCTTACATTAATGAACCTAGTGTGGGAGCTGTACCT 772
Db 812 GAAACTCAAACTGGGAAAGTGTATATATCTTGAATTTGGCTTCAGCAGCTATTTCT 871
QY 773 GGCAAGTATACATCTGTTGTTGTTGGGATTAATCTTCGAGTCTTCTCTGTGTTCTCCA 832
Db 872 GGCAAGTATACATCTGTTGTTGTTGGGATTAATCTTCGAGTCTTCTCTGTGTTCTCCA 931
QY 833 ATGCAATAAGGTTTGGGACTCCAGTGGTCTTCTATCTTGGCTGTGTAATCATTTTCCATG 892
Db 932 ATTCATAAATCTGCTGTGGGACTGCTTATAGTTCCAGTTGTAGCAGTGTATAGTTTCCATG 991
QY 893 ACAAATGAATGGTTAAAGTGTATTTCTATGATCTCTAGCTATTTGGGGTTTCACTTCTCT 952
Db 992 ATAAGATGGATGCAATCAAGATTTTCTCCATCATTTTAGCTATCTGGGGCTTCTTTCAT 1051
QY 953 ATGTCTACCAACATATCTTGTATGACAAAAAATTTGAAGA 991
Db 1052 TTGTCTATCAGCACTACTCTCGACGAAAAGATTGAAGA 1090

RESULT 14

LOCUS

AY074546 1428 bp mRNA linear PLN 18-SEP-2002
Arabidopsis thaliana unknown protein (At4g18200) mRNA, complete cds.

ACCESSION

AY074546.1

VERSION

AY074546.1

KEYWORDS

FLI CDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1428)

Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,

Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,

Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,

Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,

Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,

Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,

Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Full Length cDNA Clones

JOURNAL REFERENCE	Unpublished 2 (bases 1 to 1428)
AUTHORS	Yamada, K., Banh, J.J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheun, K., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
TITLE	Direct Submission
JOURNAL	Submitted (16-JAN-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT	Riken Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAF1 cDNAs: Yamada.K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

```

FEATURES             source
1. .1428
   /organism="Arabidopsis thaliana"
   /mol_type="mRNA"
   /db_xref="taxon:3702"
   /chromosome="4"
   /clone="RAPL09-36-N15 (R18628)"
   /ecotype="Columbia"
   /note="This clone is in a modified pBluescript vector
(F1C-1) as a BamHI/XhoI insert."

```

(F3C-17) as a *Baculini/Anoli* insect.

gene
1. .1428
/gene="At4g18200"
5'UTR
1. .71
/gene="At4g18200"
72. .1205
/gene="At4g18200"
CDS
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="AA169512.1"
/db_xref="GI:18491221"
/translation="MGISQHYHNCNQDQLEALLDHEETESFSPQPTKNCRWLRVSI
YATFVFCQPLATVILGYENGKSTVYVTLQLGFPVLILFRFPRIRQPKSDT
NFSQSPFTTASVLCGTGLLSAVYALSAVGLLYLPVSTFSLILASOLATFAFSYF
LSQKSTPTLLVNSLFTLLTVSSALLVNTDSENTNVSRYQYVIGFICTIGASAGILV
LNSQILFRKVFTHKTSASVLDLIANYQSLVATCVLIGLPSAGEWRTLPSEMRNYKLG
KSYIITLASAAALFWQYITGVCGVLIPESSVFSNSITAVGLPIVPVVAIVIFDHKMD
AKSIFGIIILAIINGFLSPVYQHYLDEKKLKTCTQPKPVEETQTLE"
1206. .1428
/gene="At4g18200"
misc_difference 1412
/gene="At4g18200"
/note="not present in genomic sequence"

3'UTR	ASAFISLLALAIWGLSFVIOHYHLDLKKKIKTCQTRPVEEETQIL"
	1206. .1428
misc_difference	1412
	/gene="At4g18200"
	/gene="At4g18200"
	/note="not present in genomic sequence"

ORIGIN

Query Match 34.1%; Score 368.2; DB 8; Length 1428;
Best Local Similarity 62.8%; Pred. No. 1.4e-79;
Matches 590; Conservative 0; Mismatches 343; Indels 6; Gaps 1;

KEYWORDS SOURCE ORGANISM

FLI CDNA; CAP trapper.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE AUTHORS

1 The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team;
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuka, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Otomo, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 1476)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akashira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yanada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and
Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Teukuba, Ibaraki
305-8602, Japan [E-mail: skikuchi@ias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007]
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., and
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., and
Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K., and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,

TITLE JOURNAL

COMMENT

FEATURES

source

1..1476
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultiivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023087G23"

ORIGIN

Query Match 28.4%; Score 307; DB 8; Length 1476;
Best Local Similarity 59.4%; Pred. No. 1.5e-64;
Matches 545; Conservative 0; Mismatches 360; Indels 12; Gaps 1;

Qy 92 TATATATGCAACGGAGGAAACAGTAATGGCTAGCAACGGTAGTTCAACTTGTGGCT 151
Db |||||
Qy 326 TCTACTACAATCAAGCGGCAATAGCAAGTGGATGCCAATTCGTCCAAACTGCTGCT 385
Db |||||
Qy 152 TTCC-----TGTGCTACTTCCATATATATATCTTGTCTATTTAAACACATGCAA 199
Db |||||
Qy 386 TTCCGATTTTGTTCATTTGCCCTATTTCTTTTCCATTCAAAGACATCTTCTACACAAACAG 445
Qy 200 CAACGTATAGAGATGGAAGAAGAACCTCAGCTAGGAACCGGTATTTGGTTAGCTAGTGC 259
Db |||||
Qy 446 TCACGTAGTCTCTGCCCTCACAATTTCTATCCCAAAATTAATCTCTGATATATGTTGTC 505
Qy 260 TTGACCTCTTGTAGGAGCAGATTGCTATCTGTACTTCCATTGGACTTCTTTTACTTACCCG 319
Db |||||
Qy 506 TGGGCCCTCATCATTTGCTGCAGACGACTTGGTGTATTTCTTATGGCTACTATATCTTCGG 565
Qy 320 TTTCTACTATTTCCCTGATCTGTGCATCTCAGTTAGCCCTTCAATGCTTTCTTCTTTATT 379
Db |||||
Qy 380 TTCTTAATCACAATAACTTACCCCTATCATTTTAAATTTCTTTTCTTCTTAACTATAT 439
Db |||||
Qy 626 TCCTCAATGCTCAAAAAATTCACCCCTCTGATTTTCAATTCGGTAGTCTCTCTTACGTTTT 685
Qy 440 CTTCCACCCTACTTTCATTAATGAGGAGCAGACTCCACAAAAGTTTCAAAAAGGAG 499
Db |||||
Qy 686 CTGCTTCACTCTTGGAGTTGATGAAGATTCTCAGGAACTACTAGTATATCATCATGGGA 745
Qy 500 AGTATGTCAAAGGTTTCATATGCAACCGTTGCTGCGTCTGCTGGTTATGGTTAGCTTTAT 559
Db |||||
Qy 746 AGTACATTTTGGGTTTCTCTGTGACACTAGGGGCATCAGCTACATCTCGCTCATTTCTCT 805
Qy 560 CCCTACAAAGCTAGTCCCTTTCTAAAGTCTTAAAGAGCAAAATTTCTCAGAAGTTATGG 619
Db |||||
Qy 806 CCCTGATGCAAGTCACATTTTGAAGGTTTATTAAGAGGAGACCTTCTCAGTTGTGTGA 865
Qy 620 ATATGATAAATCTACGTGAGTCTAGTGGCCAGTTTGTGTAGCTGGTGGGCTTTTGTGTA 679
Db |||||
Qy 866 ACATGCGATATATACAGCTCTCTGTGGCAACATTTGGCTTCTTCTTGTGGGTTATTTGCAA 925
Qy 680 GCAGTGAAGTGAAGAACTTTGAGCAGTAGTGAATGGATAACTACAAACATGGGAAGTATCCT 739
Db |||||
Qy 926 GTGGTGAATGGATGACTTTTACAAGGAGAGATGCAATTCAGTCTGGGAAGCTGTCAT 985
Qy 740 ACATTATGAACCTAGTGTGGACAGCTGTACCTGGCAGTTATTTCTCCATCGTGGCACAG 799
Db |||||
Qy 986 ATGTAATGACACTGCTGTGGACGGCTATATCTTTGGCAGGTAGCATCATAGTTGGAGTGGTGG 1045

```
Qy      800 GACTGATCTTCGAGCTCTCCTCTATCTCAAAATGCAATAGCGTTTGGGACTCCAG 859
Db      1046 GATTGATCTTTGTGGTGTCAATCGCTGTTTCAAATGTGATAGCACCTAGCTCTACCCA 1105
Qy      860 TGGTTCTATCTTGGGCTGTAATCATTTCCATGACAAATGAATGGTTAAAGGTGATTT 919
Db      1106 TCAATCTGTTTTTGTGTGATTTCTTTCATGACAAGATGGATGGAGTAAAGATTATAG 1165
Qy      920 CTATGATCCTAGCTATTTGGGGTTTCACCTTCTCTATGTTACCAACAATATCTTGTATGACA 979
Db      1166 CTATGCTGATGGCCATTTGGGGATTATGTCATATGCCCAATATATGTTGTATGGCA 1225
Qy      980 AAACTTGAAGAAAAT 996
Db      1226 AGAAGGTAGAAAGACT 1242
```

Search completed: November 1, 2004, 15:44:48
Job time : 4796.3 secs

THIS PAGE BLANK (USP 107)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 02:40:28 ; Search time 560.296 Seconds
(without alignments)
9996.736 Million cell updates/sec

Title: US-09-913-767-10

Perfect score: 1067
Sequence: 1 agacaagaatgtgaaggct.....gaccataagtgaagttaa 1067

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	476.4	44.6	1049	3	AAA97920 A. thalia
2	428.4	40.1	1225	3	AAA97919 A. thalia
3	342.8	32.1	2586	3	AAC44184 Arabidops
4	177.4	16.6	493	3	AAC36831 Arabidops
5	147	13.8	1071	3	AAA97925 A. thalia
6	145.6	13.6	3387	6	ABZ14407 Arabidops
7	145.6	13.6	3387	8	ADA68246 Arabidops
8	141.6	13.3	1152	8	ADA70213 Rice gene
9	136.2	12.8	1194	3	AAA97923 A. thalia
10	135	12.7	1047	6	ADG88190 A. thalia
11	135	12.7	1081	3	AAA97924 A. thalia
12	131.4	12.3	1200	8	ADA70627 Rice gene
13	119.8	11.2	1293	3	AAA97922 A. thalia
14	116.6	10.9	1145	3	AAA97921 A. thalia
15	116.6	10.9	1292	3	AAC51119 Arabidops
16	116.6	10.9	1295	3	AAC37289 Arabidops
17	114	10.7	83698	6	ABN85767 Arabidops
18	112.8	10.6	2660	3	AAC47974 Arabidops
19	112.6	10.6	1403	3	AAC37652 Arabidops
20	107.6	10.1	1418	3	AAC46395 Arabidops
21	107.6	10.1	1421	3	AAC39779 Arabidops

22	81.6	7.6	2175	8	ADA70748	Ada70748 Rice gene
23	72	6.7	592	10	ABX56692	Abx56692 Arabidops
24	51.6	4.8	1155	8	ADA69686	Ada69686 Rice gene
25	49.8	4.7	2000	8	ADA71938	Ada71938 Rice gene
26	47	4.4	774	6	ABN99127	ABn99127 Arabidops
27	46.8	4.4	477	3	AAC37273	Aac37273 Arabidops
28	42.2	4.0	41104	6	AAD36260	Aad36260 Human G-p
29	41	3.8	1040	8	ADA69568	Ada69568 Rice gene
30	40.6	3.8	6317	6	ABL32408	Ab132408 Human inm
31	40.6	3.8	6317	6	ABL49311	Ab149311 Human pol
32	39.4	3.7	9974	2	AAX20573	Aax20573 Polynucle
33	39.2	3.7	478	9	ACH35469	Ach35469 Human end
34	39.2	3.7	896	8	ABX05058	Abx05058 Human nov
35	39	3.7	110000	6	ABN71527_06	Continuation (7 of
36	38	3.6	2000	8	ADA71938	Ada71938 Rice gene
37	38	3.6	6621	6	ABK33933	Abk33933 Human DNA
38	38	3.6	6621	6	ABL70156	Ab170156 Chemical
39	38	3.6	6621	6	AAS61098	Aas61098 Human gen
40	38	3.6	6621	8	ABZ09984	Abz09984 Haematopo
41	38	3.6	6621	8	ABZ10098	Abz10098 Haematopo
42	38	3.6	6621	8	ADA20345	Ada20345 Prostate
43	38	3.6	6621	8	ADA84152	Ada84152 Human ren
44	38	3.6	6621	10	ADE84078	Ade84078 Human lym
45	37.6	3.5	1024	6	ABX65613	Abx65613 Helicobac

ALIGNMENTS

RESULT 1
AAA97920
ID AAA97920 standard; DNA; 1049 BP.
AC AAA97920;
XX
DT 19-JAN-2001 (first entry)
XX
DE A. thaliana PUP1 DNA #2.
XX
KW PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.
XX
OS Arabidopsis thaliana.
XX
PN DE19907209-A1.
XX
PD 24-AUG-2000.
XX
PF 19-FEB-1999; 99DE-01007209.
XX
PR 19-FEB-1999; 99DE-01007209.
XX
(FROM/) FROMMER W.
XX
Gillissen B, Buerkle L, Andre B, Frommer WB;
XX
WPI; 2000-566202/53.
XX
Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transporter, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.
XX
Claim 2; Page 12; 24pp; German.
XX

This invention describes a novel nucleic acid encoding a plant nucleobase transporter (I). (I) is produced by complementation of a nucleobase transport (NET)-defective host cell with a plant gene bank by selection of NET-positive cells. (I) is used to isolate homologous sequences from bacteria, fungi, plants, animals and humans, for expression of the encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting expression of (II) (when in antisense orientation), and to produce transgenic crop plants. The transgenic plants have modified nucleobase transport properties, e.g. altered affinity and substrate specificity

PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140699P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142059P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145911P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
			Query Match 32.1%; Score 342.8; DB 3; Length 2586;		
			Best Local Similarity 67.9%; Pred. No. 8.6e-89;		
			Matches 509; Conservative 1; Mismatches 234; Indels 6; Gaps 2;		
Qy	245	TCCTATCGCCGCTGTTATTGTAGGCATTCCTCTCAGGCTTTTGTAACTACTTGTATGCATA	304		
Db	1643	TCATATCGCTCCATTGTCATAGGGTTCCTCAGAGGACTTGACAACTACTATATCTTA	1702		
Qy	305	TGGTATAGCTATCTTCCAGTTTCTACAGTGTCTTATCATTCGCTTCACAGTTAGCTTT	364		
Db	1703	CGGATTAGCATATCTGCCAGTTTCACTTCGCTCATATCGGAACTCAACTAGCTTT	1762		
Qy	365	TATAGCTATCTTCTCATCTTTCATGGTTAAACATAAAGTTCACCTCTTTTACCATCAATGC	424		
Db	1763	CAAGCTCTCTTCGCTTCTTGTAGTCAAGCAAAAGTTTCACTCCGTTCTCCATAACGC	1822		
Qy	425	TGTTGTTGTTGACTGTTGGTGCTCGGTTTTTGGGAATGCATACCGAAACTGATAAGCC	484		
Db	1823	CGTCGTTTTTGTGACGGTTGGTATCGGGATCCCTTCGCTTACACAGTGCATGGAGACAAACC	1882		
Qy	485	AGTTCATGAGACTCACAAAGCAGTACATAAATGGTTTCTTCTGTTACTTACTAGCAGCTGT	544		
Db	1883	GGCTAAGGAGAGCAAGAAAGAGTATGTTGGGTTCTTGTATGACTGTGGTTGCAGCTCT	1942		
Qy	545	TATGTATGCTTTCATCTTGCCATTAGTGGAACTTGCTTACCAGAAAGCTAAGCAACCAT	604		

Db	1943	TCTCTATGCTTTTATATATACCGTCTGTTGAGCTAACTTACAAGAAAGCTCGTCAAGAAAT	2002	PR	06-MAY-1999;	99US-0132486P.
Qy	605	GAGCTATACCTTGTGTCGAGTTCCAGTTGATTTTGTGTCTCTCTTCTTCTATTCTGTCAG	664	PR	06-MAY-1999;	99US-0132487P.
Db	2003	CACITTTCCCACTTGTGCTTGAGATTGATGGTCATGCTGCTGCTGCTACITTTTCTG	2062	PR	11-MAY-1999;	99US-0132863P.
Qy	665	CGTCATCGGTATGTTTCATCGCTGGTGATTTTCAAGCAGGCGCTTACCAAAGAAAGCAAGAGA	724	PR	11-MAY-1999;	99US-0134256P.
Db	2063	TGTCATTTGGCATGTTTCATCGTTGGAGATTTTA---AGGTGATGCAVGAAGCAAGAGA	2119	PR	14-MAY-1999;	99US-0134218P.
Qy	725	GTTCAAGCTTGGAGAGGCATTGTT---CTATGTTGGTGGCTGTGTTTTTCAGCCCATCATATG	781	PR	14-MAY-1999;	99US-0134219P.
Db	2120	GTTCAAGATTGGAGGATCAGTGTGTTTACTATGCAATTGATAGTGATCACAGGAATAATATG	2179	PR	14-MAY-1999;	99US-0134221P.
Qy	782	GCAAGGCTTCTCTGGAGCCATTGGATTAATCTTCTCCACATGCTCTGCTCTCGCGG	841	PR	14-MAY-1999;	99US-0134370P.
Db	2180	GCAAGGTTTCTCTTAGGAGCCATAGGATGTGTTTTGTGCATCATCACTAGTCTCTGG	2239	PR	18-MAY-1999;	99US-0134768P.
Qy	842	TATTATGATATCAGTCTGTTTTTGCCAAATTACAGAGGTTTTTAGCTGTGTTATATTTACCATGA	901	PR	19-MAY-1999;	99US-0134941P.
Db	2240	TGTTCTGATAAGTGTCTGCTCCGGTGACTGAAGTTTTTCGCGTCTGTTTTCGGGA	2299	PR	20-MAY-1999;	99US-0135124P.
Qy	902	AAAGTTTCAAGCTGAGAGGACCTTCTCTGCTCTCTCCCTTTGGGGCTTTGTCTCTCTTA	961	PR	20-MAY-1999;	99US-0135353P.
Db	2300	GAAGTTTTCAGSCAGAGAAAGTGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTA	2359	PR	24-MAY-1999;	99US-0135629P.
Qy	962	CTTTTATGTTGACATAAAGTCTCGCGGGA	991	PR	25-MAY-1999;	99US-0136021P.
Db	2360	CTTCTACGGCGAGTTTAAATCCGGCAAGAA	2389	PR	27-MAY-1999;	99US-0136392P.
RESULT 4						
AAC36831						
ID	AAC36831 standard; DNA; 493 BP.					
XX	AAC36831;					
AC	AC					
XX	17-OCT-2000 (first entry)					
DT	DT					
XX	XX					
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 15222.					
KW	Hybridisation assay; genetic mapping; gene expression control;					
KW	protein identification; signal transduction pathway; metabolic pathway;					
KW	promoter; termination sequence; ss.					
XX	XX					
OS	Arabidopsis thaliana.					
XX	XX					
PN	EP1033405-A2.					
XX	XX					
PD	06-SEP-2000.					
XX	XX					
PF	25-FEB-2000; 2000EP-00301439.					
XX	XX					
PR	25-FEB-1999;	99US-0121825P.		PR	06-MAY-1999;	99US-0132486P.
PR	05-MAR-1999;	99US-0123180P.		PR	06-MAY-1999;	99US-0132487P.
PR	09-MAR-1999;	99US-0123548P.		PR	11-MAY-1999;	99US-0132863P.
PR	23-MAR-1999;	99US-0125788P.		PR	11-MAY-1999;	99US-0134256P.
PR	25-MAR-1999;	99US-0126264P.		PR	14-MAY-1999;	99US-0134218P.
PR	29-MAR-1999;	99US-0126785P.		PR	14-MAY-1999;	99US-0134219P.
PR	01-APR-1999;	99US-0127462P.		PR	14-MAY-1999;	99US-0134221P.
PR	06-APR-1999;	99US-0128234P.		PR	18-MAY-1999;	99US-0134370P.
PR	08-APR-1999;	99US-0128714P.		PR	18-MAY-1999;	99US-0134768P.
PR	16-APR-1999;	99US-0129845P.		PR	19-MAY-1999;	99US-0134941P.
PR	19-APR-1999;	99US-0130077P.		PR	20-MAY-1999;	99US-0135124P.
PR	21-APR-1999;	99US-0130449P.		PR	20-MAY-1999;	99US-0135353P.
PR	23-APR-1999;	99US-0130510P.		PR	24-MAY-1999;	99US-0135629P.
PR	28-APR-1999;	99US-0130891P.		PR	25-MAY-1999;	99US-0136021P.
PR	30-APR-1999;	99US-0131449P.		PR	27-MAY-1999;	99US-0136392P.
PR	03-MAY-1999;	99US-0132048P.		PR	28-MAY-1999;	99US-0136782P.
PR	04-MAY-1999;	99US-01322407P.		PR	01-JUN-1999;	99US-0137222P.
PR	05-MAY-1999;	99US-0132484P.		PR	03-JUN-1999;	99US-0137528P.
PR		99US-0132485P.		PR	04-JUN-1999;	99US-0137502P.

PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 25-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0145386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161992P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0162142P.
PR 04-AUG-1999;	99US-0147302P.		
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 20-AUG-1999;	99US-0149929P.		
PR 23-AUG-1999;	99US-0149802P.		
PR 23-AUG-1999;	99US-0149930P.		
PR 25-AUG-1999;	99US-0150566P.		
PR 26-AUG-1999;	99US-0150884P.		
PR 27-AUG-1999;	99US-0151065P.		
PR 27-AUG-1999;	99US-0151066P.		
PR 27-AUG-1999;	99US-0151080P.		
PR 30-AUG-1999;	99US-0151303P.		
PR 31-AUG-1999;	99US-0151438P.		
PR 01-SEP-1999;	99US-0151930P.		
PR 07-SEP-1999;	99US-0152363P.		
PR 10-SEP-1999;	99US-0153070P.		
PR 13-SEP-1999;	99US-0153758P.		
PR 15-SEP-1999;	99US-0154018P.		
PR 16-SEP-1999;	99US-0154039P.		
PR 20-SEP-1999;	99US-0154779P.		
PR 22-SEP-1999;	99US-0155139P.		
PR 23-SEP-1999;	99US-0155486P.		
PR 24-SEP-1999;	99US-0157865P.		
PR 28-SEP-1999;	99US-0158029P.		
PR 29-SEP-1999;	99US-0158232P.		
PR 04-OCT-1999;	99US-0158369P.		
PR 13-OCT-1999;	99US-0159293P.		
PR 13-OCT-1999;	99US-0159294P.		
PR 13-OCT-1999;	99US-0159295P.		
PR 14-OCT-1999;	99US-0159329P.		
PR 14-OCT-1999;	99US-0159330P.		
PR 14-OCT-1999;	99US-0159331P.		
PR 14-OCT-1999;	99US-0159637P.		
PR 18-OCT-1999;	99US-0159638P.		
PR 21-OCT-1999;	99US-0159584P.		
PR 21-OCT-1999;	99US-0160741P.		
PR 21-OCT-1999;	99US-0160767P.		
PR 21-OCT-1999;	99US-0160768P.		
PR 21-OCT-1999;	99US-0160770P.		
Query Match 16.6%; Score 177.4; DB 3; Length 493;			
Best Local Similarity 65.1%; Pred. No. 5.6e-41;			
Matches 287; Conservative 0; Mismatches 136; Indels 18; Gaps 1;			
Qy 1	AGACAAGATGGTGAAGGCTCTTGTGATCATATAAAGTGCATATAATCTAGCCATAGGAAACT 60		
Db 53	AGAGAAGATGAAGAATGGTTTGATAATCATATAAACTGTATTATCTCACTATAGGAACAT 112		
Qy 61	GTGGAGGTCCCTTTGATTATGCGTCTTACTTCAACAATGGCGTAAAGGATTTGGTTCT 120		
Db 113	GTGGAGGTCCCTTTGTTAACTGCTCTTACTTCCAAATGGCGGAAACGAATCTGGTTCA 172		
Qy 121	CTACGTTTCTTGAAGTGCAGGCTTTCCTGTATCTTCAATCTCTCTCTCTCTCTCTTACA 180		
Db 173	TGAGCTTCTTATCAACCGCTGGTTTTCATCATCTCTCTCTCTCTCTCTCTCTCTCTCC 232		
Qy 181	TTACCGCGCGAGCAACAATGTGGTGA-----TAGTACAAGTT 222		
Db 233	TCAGCGTCGCGCAGCAACCGCAACCCCTAACACGCGGAAACAAAGCGGAAACAAAGC 292		
Qy 223	TCCTTTCTTATCAACCGCTCTTCTTATCGCGCTGTTATTGTAGGCATTTCTCTCAGGGT 282		
Db 293	TCCTTCTCATGGAACCTCTCTGTTTATCGCTCCATGTCATAGGGTTGCTCACAGGAC 352		
Qy 283	TTGATAACTACTTGTATGCATATAGGTATAGCTTATCTTCCAGTTTCTACAGCTCTCTTA 342		
Db 353	TTGACAACTACTTATATTCTTACGATTAGCATATCTGCCAGTTTCAACTTCACTCGCTCA 412		
Qy 343	TCATTGCTTCACAGTTAGCTTTTATAGCTATCTTCTCATTTCTTCAATGTTTAACATAAGT 402		
Db 413	TAATCGGAACCTCAACTAGCTTTCACGCTCTCTTTCGCTTTCTTGTAGTCAAGCAAAAGT 472		
Qy 403	TCACTCCTTTTACCATCAATG 423		
Db 473	TCACTCGTTCTCCATAAAG 493		
RESULT 5			
AAA97925			
ID	AAA97925 standard; DNA; 1071 BP.		
XX	AAA97925;		
AC	AAA97925;		
XX	19-JAN-2001 (first entry)		
DT	A. thaliana PUP1 DNA #7.		
XX	PUP1; transgenic plant; nucleobase transporter; apical dominance;		
XX	flowering behaviour; senescence; pesticide distribution; ds.		
XX	Arabidopsis thaliana.		
XX	DE19907209-Al.		
XX	24-AUG-2000.		
XX	19-FEB-1999; 99DE-01007209.		

XX 19-FEB-1999; 99DE-01007209.
XX (FROM/) FROMMER W.
XX Gillissen B, Buerkle L, Andre B, Frommer WB;
XX WPI; 2000-566202/53.
XX Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transport, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.
XX Claim 2; Page 15; 24pp; German.
XX This invention describes a novel nucleic acid encoding a plant nucleobase
CC transporter (I). (I) is produced by complementation of a nucleobase
CC transport (NBT)-defective host cell with a plant gene bank by selection
CC of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
CC expression of (II) (when in antisense orientation), and to produce
CC transgenic crop plants. The transgenic plants have modified nucleobase
CC transport properties, e.g. altered affinity and substrate specificity
CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX
SQ Sequence 1071 BP; 307 A; 236 C; 208 G; 320 T; 0 U; 0 Other;
Query Match 13.8%; Score 147; DB 3; Length 1071;
Best Local Similarity 49.8%; Pred. No. 5.7e-32;
Matches 462; Conservative 0; Mismatches 450; Indels 15; Gaps 3;
QY 41 AATTCTAGCCATAGGAAGTGGAGTCTTGGATTATGCTCTCTACTTCAACAATGG 100
DB 39 AGTCAAGTCATGGCCATCATGCTGCTACAAATCTGGGCAGACTATATGAAATGG 98
QY 101 CGGTAAAGGATTTGGTCTCTACGTTTCTCAAACTGCAGGCTTCTCTGTTATCTTCAT 160
DB 99 AGGAACAGCAATGGCTAGCAACGGTAGTTCAGCTTGTAGGCTTCTCTATCTACTTCC 158
QY 161 TCCTCTGCTCTCTCTTACATTTACCCGGCGAGAGCAACAATGGGTAGTACAAAG 220
DB 159 ATATCATCTCTTGTGCTAGTCAAAACACATACAACTCAGA-----GAGATGGCAATT 212
QY 221 TTTCTTTCTTATCAAAACCGGCTCTTATTCGCGCGTGTATTGAGGATTTCTCTCAGG 280
DB 213 AACCTCACTTAGGAACCGGTGCAATTAGTTTAC-----ATAGTGTCTGGACTTCTTGTAGG 266
QY 281 GTTTCATTAACACTTGTATGATATGGTATAGCTTATCTTCCAGTTTCTACAGCTGCTCT 340
DB 267 AGCAGCTGTCTACCTATATTCATTTGAGTGTCTTACCTACCTGTTTCTACCTTTCCCT 326
QY 341 TATCATTTGCTTACAGTGTAGCTTTTATAGTATCTTCTCATTTCTTCTGTTTAAACATAA 400
DB 327 GATCTGTGCATCACAGTTAGCTTCCAGCTTTTCTTCTTATTTACTCAACTCAACAAA 386
QY 401 GTTCACTCTTTTACATCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGG 460
DB 387 ACTTACTCTCATTTTGAATTTCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 446
QY 461 AATGATACCGAACTGATTAAGCCAGTTTCATGAGTCTACAGAGTACATCACTGGTTT 520
DB 447 ATTTAATACAGAGGAATCAGATTCACAAAAGTTACAAAGAGAGTATGTCAAAGGTTT 506
QY 521 CTTGATTTACTGTAGCAGCAGCTGTTTATGTATGCTTTTCTTCTTCTTCTTCTTCTTCTG 580
DB 507 CGTATGCACCGTTGTGTCATCTGCTGGGTTTGGTCTACTCTTATCCCTACACACAGTAGC 566
QY 581 TTACAGAAAGCTTAAGCAAAACCATGAGCTATACCTTGTGCTGAGTTCAGTTGATTTT 640

DB 567 CTTTCGTAAGTTTTTAAAGAACAACTTTCTCAGAGTTTAAATATGATAATCTACAT 626
QY 641 GTGCTCTCTGCTTCTTATTTGTGAGGTCATCGGTATGTTTCGCTGGTGTATTTCAAGCA 700
DB 627 GAGTCTAGTGGCCAGTTGTGTAGGCTGGTGGG---GCTTTTGTAGTAGGAGTGGAA 683
QY 701 GGCTTTACCAAAAGCAAGAGAGATTCAAGCTTGGAGAGGAGTGTCTTATGTGGTGGC 760
DB 684 AACTTTGAGCAGTGAATGGAAGAACTACAAACTTTGGGAAGGTATCTCTATGTATGAACCT 743
QY 761 TGTGTTTTTCCAGCATCATATGCAAGGCTTCTTCTGGAGGCCATTTGATTAATCTTCTC 820
DB 744 AGTGTGACAGCTGTACTGCGAGGTATTTCTCCATCGGTTTCCAGGACTGATCTTGA 803
QY 821 CACATCGTCTCTGCTCGGGTATTTATGATATCATGATGCTTTTGGCAATTCACAGAGTTT 880
DB 804 GCTTCTCTCCCTATTTCTCAATGCAATTAAGCGCTTGGGACTCCCGTGGTTCCTATCCT 863
QY 881 AGCTGTATATTTACCATGAAAGATTTCAGAGTGTGAGAGGAGCTTTCTCTTGTCTCTC 940
DB 864 GGTCTCATCATTTTCCATGACAAATGAACGGCTTAAAGGTGATTTCTATGATTTCTAGC 923
QY 941 CCTTTGGGGCTTTGCTCTTACTTTTA 967
DB 924 TATTTGGGGTTTCGTATCCTATGCTA 950
RESULT 6
ABZ14407
ID ABZ14407 standard; DNA; 3387 BP.
XX
AC ABZ14407;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2212.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; da.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PP 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Krepes J, Wang X, Zhu T;
XX
WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 2212; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid with an
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is

CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office

Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;

Query Match	13.6%	Score 145.6	DB 6	Length 3387
Best Local Similarity	49.1%	Pred. No. 2.5e-31		
Matches 475	Conservative 0	Mismatches 484	Indels 9	Gaps 3
Qy	1	AGACAAGATGGTGAAGGCTCTTGATCATAAACTCATAAATCTAGCCATAGGAACCT	60	
Db	2354	AGAACTGAAGGTGGCTCGGTCTCCATATACGAATCTTTGTCATCTTCTGCCAAC	2413	
Qy	61	GTGGAGTCCCTTGAATATGCGTCTACTTCAACAATGGCGGTAAAGGATTTGGTTCT	120	
Db	2414	CACCTGCTACAGTCTCGGTAGACTGTACTATGAAATGGAGGAAAGCACATATGTGG	2473	
Qy	121	CTAGCTTCTTGAAGTGCAGGCTTTCCTGTTATCTTCATCTCTCTGCTCTCTCTTACA	180	
Db	2474	TAACACTTCTTCAACTCATTTGGCTTCCTGTACTGATTTCTTCCGCTCTTTTCTCGAA	2533	
Qy	181	TTACCCGGCGCAGAAGCAACAATGTGGGTGATAGTACAAGTTTCTTTCTTATCAACCGC	240	
Db	2534	TCAGGCAACCAATCAACAGATACAATTTCACT--CAGTCCCTTCTTCAACCCT	2591	
Qy	241	GTCTTCTTATGCGCGCTGTATTGTAGGCAATTTCTCAGGCTTTGATPAACTACTGTATG	300	
Db	2592	TGCATCGGTTACTTGTGCACT----GGACTGCTAGTGTCTGCTTATGCTTATTTGTCTG	2647	
Qy	301	CATATGATAGCTTATCTTCCAGTTTCTACAGCTGCTCTTATCATTTGCTTCACAGTAG	360	
Db	2648	CAGTTGGGTGCTTTACTTACCAGTCTCTACTTTCTCCCTCATCTTGGCCTCACAGTTG	2707	
Qy	361	CTTTTATAGCTATCTTCTCATCTTCTATGGTTAAACATAAAGTTCACCTCTTTTACCATCA	420	
Db	2708	CCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAGTTCACCTCTTGTATGATCA	2767	
Qy	421	ATGCTGTTGTGTTTGACTGTTGGTGCTGCGGTTTTGGGAAATGCATACCGAAACTGATA	480	
Db	2768	ATTCCTTTGTTTCTCCTTACGGTTTTCTCTGCCCCCTCTCGTGGTCAACACTGATTCAGAAA	2827	
Qy	481	AGCCAGTTCTAGAGACTCACAGCAGTACATAACTGCTTTCTTGATTTACTGTAGCAGCAG	540	
Db	2828	ACAACAATATGATATAGAGTACAATGTGATTTGGGTTTCACTGTACCAATTTGGTCTT	2887	
Qy	541	CTGTTATGATGCTTTTCATCTTGCCATTAGTGGAACTTGCTTACGAGAAAGCTAAGCAAA	600	
Db	2888	CCGCTGGATTGGATTGGTACTACTCTGTATACAACTGCTCTTCAGGAAGTTTTCAGGA	2947	
Qy	601	CCATGAGCTATACCCCTGTGCTCGAGTTTCAGTTGATTTTGTGTCCTCTCTGCTTCTATTG	660	
Db	2948	AGCATACATCTCAGCAGTCTGAGCTTTGGCCAAATTACCAGTCTCTAGTTGCAACTTGTG	3007	
Qy	661	TCAGCGTATCGGTATGTTTCATCGCTGGTGAATTTCAAGCAGGCTTACCAAAAGAACAA	720	
Db	3008	TGGTACTCATAGGACTGTTTGCAGAGTGGAGT---GGAGAACTCTGCCAAGTGAATGA	3064	
Qy	721	GAGAGTTCAAGCTTGGAGGCAATGTTTCTATGTGGTGGCTGTGTTTTCAGCCATCATAT	780	
Db	3065	GAAACTACAACTGGGAAAGTGTACATATCTTTGACTTTTGGCCTCAGCAGCTATTTTCT	3124	
Qy	781	GGCAAGGCTTCTTCGGGACCAATGGATTAATCTTCTCCACATCGTCTCTCGTCTCGG	840	
Db	3125	GGCAAGTATACATGTTTGGTTGTGTGGGATTAATCTTCAGTCTCTTCTCTGTGTTCTCCA	3184	
Qy	841	GTATTATGATACAGTCTTTTGGCAATTACAGAGGTTTTAGCTGTTTATATTCTACCATG	900	
Db	3185	ATTCATTAACCTGTGGGACTGCTATAGTTCCAGTTGTAGCAGTGAATGTTTCCATG	3244	
Qy	901	AAAAGTTTCAAGCTGAGAAGGACTTTTCTCTGTCTCTCTCCCTTTGGGCTTTGTCTCTT	960	
Db	3245	ATAAGATGGATGATCAAGAAATTTTCCATCATTTTGTAGTATCTGGGGCTTCTTTTCAT	3304	

Db 847 GAGACCTTTCTCAGTTGTGTGAACATGCAGATATATACAGCTCTCGTGGCAACATTGGCT 906
QY 663 AGCGTCATCGGTATGTTTCATCGCTGGTGAATTCAAG 698
Db 907 TCTCTTGTGGGTATTTTGAAGTGGTGAATGGATG 942

RESULT 9
AAA97923
ID AAA97923 standard; DNA; 1194 BP.
AC AAA97923;
XX
XX 19-JAN-2001 (first entry)
XX
XX A. thaliana PUP1 DNA #5.
XX
XX PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.
XX
XX Arabidopsis thaliana.
XX
XX DE19907209-A1.
XX
XX PD 24-AUG-2000.
XX
XX PF 19-FEB-1999; 99DE-01007209.
XX
XX PR 19-FEB-1999; 99DE-01007209.
XX
XX (FROM/) FROMMER W.
XX
XX PA Gillissen B, Buerkle L, Andre B, Frommer WB;
XX
XX DR WPI; 2000-566202/53.
XX
XX PT Nucleic acid, useful for producing transgenic plants with altered
PT nucleobase transport, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.
XX
XX PS Claim 1f; Page 14; 24pp; German.

XX This invention describes a novel nucleic acid encoding a plant nucleobase
CC transporter (I). (I) is produced by complementation of a nucleobase
CC transport (NBT)-defective host cell with a plant gene bank by selection
CC of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting
CC expression of (II) (when in antisense orientation), and to produce
CC transgenic crop plants. The transgenic plants have modified nucleobase
CC transport properties, e.g. altered affinity and substrate specificity
CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention
XX

XX Sequence 1194 BP; 298 A; 279 C; 231 G; 386 T; 0 U; 0 Other;
SQ

Query Match 12.8%; Score 136.2; DB 3; Length 1194;
Best Local Similarity 48.7%; Pred. No. 8.4e-29;
Matches 462; Conservative 0; Mismatches 478; Indels 9; Gaps 3;

QY 20 TCTTGTGATCATAACTGCATTAATCTAGCCATAGGAAACTGTGGAGTCCTTTGATTAT 79
Db 135 TCGTGTCTCCCATATACGPAATCTTTGTCTCTTTTGGCAGCCACTAGTACAAATCTGGG 194
QY 80 CGGTCTCTACTTCAACAATGGCGGTAAAGGATTGGTCTCTAGCTTTCTTGAAGACTGC 139
Db 195 TAGATTGTACTATGAATATGAGGAAATAGACATATGTGTAACACTTCTTCACTCAT 254
QY 140 AGGCTTTCCTGTTATCTTTCATTCCTCTCTCTTTCATATTACCCGGCCGAGAACAA 199
Db 255 TGGCTTCCCTGACTGTTCTGTTCGGCTTCTTTTCTCGAATCAGGCAACCCAAATCAAC 314

QY 200 CAATGTGGTGATAGTACAAAGTTTCTTTCTTATCAAAACCGGCTCTTCTTATCGCGGCTGT 259
Db 315 AGATACAAATTTTCAGT--CAGTCCCCTTCCCTTACCACCCCTTGCATCGGTACTTGTGTC 372
QY 260 TATTGTAGGCATCTCTCAGGGTTTGTAACTACTTCTATGCATATAGCATATAGCTTATCT 319
Db 373 ACT---GGACTGCTAGTGTCCGCTTATGCTTATTGTCTGCAGTAGGGTGTCTACTT 428
QY 320 TCCAGTTTCTACAGCTGCTCTTATCATTTGTTTCAAGTTAGCTTTTATAGCTATCTTCTC 379
Db 429 ACCAGTCTCTACTTTTCTCCCTCATCTTGGCTCAGAGTTGSCCTTCACTGCTTTTCTC 488
QY 380 ATTCTTCATGTTAAACATAAAGTTCACTCTTTTACCATCAATGCTCTGTTGTTGTGAC 439
Db 489 ATATTTCCTTAACTCGCAAAAGTTCACTCTCTTTGATAGTCAGTTCTTTTGTCTCTCAC 548
QY 440 TGTGTGCTGCTCGGTTTGGGAATGCATACGAAACTGATAAGCCAGTTTCATGAGACTCA 499
Db 549 TGATCTCTGCTCTTCTTGTGTCAACACTGATTGAGAAAACTCAACTAATGATCTAG 608
QY 500 CAAGCAGTACATAACTGTTTCTTGTATTACTGTAGCAGCAGCTGTTATGTAAGTCTTCAT 559
Db 609 AGTACAGTATGTCGTTTCCATATGATACCATCGGTGCTTCCGCTGGGATTGACTGTT 668
QY 560 CTTGCCATTAGTGAACCTTGTACCAAGAAAGCTAAGCAACCATGAGCTATACCTTGT 619
Db 669 ACTATCTCTGATACAAATGCTCTTCAGGAAAGTTTTCACGAAGCATACATCCTCAGCAGT 728
QY 620 GCTCAGTTCCAGTTGATTTTGTGCTCTCTTGTCTTCTATTGTGTCAGCGTCATCGGTATGTT 679
Db 729 CACGCACTTGGCCATTTTACCAGTCTCTAGTTGCGAGTTGTAGTTCTCATAGCACTTT 788
QY 680 CATCGTGGTGAATTTCAAGCAGCGCTTACCAAAAGAAAGAGAGAGTTCAAGCTTGGAGA 739
Db 789 TGCAAGTGGAGAGT---GGGAAACTTTGCCAAGTGAGATGAGAAACTACAAACTCGGGAA 845
QY 740 GGCATGTTCTTATGTGCTGCTGTTTTCAGCCATCATATGCAAGGCTCTTCTTCTGGG 799
Db 846 AGTGTATATGTTTGTAGCTTTCAGCTCGGAGCTATTTCTCTGCAAGTCTACACTTTGG 905
QY 800 AGCCATTGGATTAATCTTCTCCATCGTCTCTCGGCTCGGGTATTATGATATCAGTGTCT 859
Db 906 TCTTGTGGGATGATCTTCGAGTCATCTCTGTGTTCTCCAAATTCATACAGCTGTGGG 965
QY 860 TTTGCCAAATTACAGAGGTTTGTAGCTGTATATTTCTACATGAAAGTTTCAAGCTGAGAA 919
Db 966 ATTGCTTATAGTTCCAGTTTGGGCGAGTGATAGTTTTCATCATAGATAAGACGCATCAAA 1025
QY 920 GGGACTTTCTCTGCTCTCTCCCTTTGGGGCTTTGTCTCTTACTTTTAT 968
Db 1026 AATCTTCCATATTATTAGCTATCTCGGGCTTCTCTTTCATTCGTCTAT 1074

RESULT 10

ADG88190
ID ADG88190 standard; cDNA; 1047 BP.

XX AC ADG88190;

XX AC 22-APR-2004 (first entry)

XX DE A. thaliana RPP4-upregulated pathogen infection-related gene #632.
XX
XX Pathogen infection-related gene; plant; Peronospora parasitica;
KW defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;
KW fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX Arabidopsis thaliana.

XX WO200222675-A2.

XX PD 21-MAR-2002.

Db 707 TAAATATGTATAGGATATATATGACATTTGGTCTTCTGCTGGGATTCGATTCGTC 766
Qy 562 TGCATATAGTGAACCTCTACCAAGAAAGTAAAGCAACCATGAGCTATACCTTTGTC 621
Db 767 TATCCCTGGTACAACTGATCCTCAGGAAGGTTTAAAGAAAGCAAACTTCTCAACGGTCA 826
Qy 622 TCGAGTCCAGTGTATTTTGTCTCCCTGCTTCTTATTTGTGAGCGTCACTCGTATGTCA 681
Db 827 CTGACTTGGTTCGCTTACCAATCTCTAGTTGCAAGCTGTGTCTCATAGGACTTTTCG 886
Qy 682 TCGCTGGTGTATTTCAAGCAGGCTTTACCAAAAGCAAGAGATTCAAGCTTGGAGAGG 741
Db 887 CAACGGGGAGT--GGAATACTTTAAACAAGTGAATGGAATACTACAACTGGGAAAG 943
Qy 742 CATTTGTTCTATGTGTGCTGTGTTTTTCAGGCATCATATGCAAGGCTTCTTTGGAG 801
Db 944 TGCCATAGCTTATGACTTTGGCTCGATAGCTATTTCTCGCAAGTCTACACCATTTGGC 1003
Qy 802 CCATTGGATTAATCTTCTCCATCGTCTCTCGTCTCGGTTATTTATGATATCAGTGTCTT 861
Db 1004 TCGTGGGACTGATCTTTGAGTCATCTCTGTGTTCTCCAAATTCATTAAGTCTGTGGAT 1063
Qy 862 TGCCAAATACAGAGGTTTTAGCTGTATATATCTACCATGAAAGTTTCAAGCTCGAAGG 921
Db 1064 TGCCTATAGTTCCAGTTGTAGCAGTGTATTTTCCATGATAAATGAACGCTCAAGA 1123
Qy 922 GACTTCTCTGCTCTCTCCCTTTGGGCTTTGCTCTTACTTTTAT 968
Db 1124 TCTTCTCCATCATTTTAGCTATCTGGGATTCATTTCAATTTGCTAT 1170

RESULT 14

AAA97921
ID AAA97921 standard; DNA; 1145 BP.

XX AC AAA97921;

XX DT 19-JAN-2001 (first entry)

XX DE A. thaliana PUP1 DNA #3.

XX PU PUP1; transgenic plant; nucleobase transporter; apical dominance;
KW flowering behaviour; senescence; pesticide distribution; ds.

XX OS Arabidopsis thaliana.

XX PN DE19907209-A1.

XX PD 24-AUG-2000.

XX PF 19-FEB-1999; 99DE-01007209.

XX PR 19-FEB-1999; 99DE-01007209.

XX PA (FROM/) FROMMER W.

XX PI Gillissen B, Buerkle L, Andre B, Frommer WB;

XX DR WPI; 2000-566202/53.

PT Nucleic acid, useful for producing transgenic plants with altered
FT nucleobase transport, encodes a nucleobase transporter protein of
PT Arabidopsis thaliana.

XX PS Claim 1f; Page 12-13; 24pp; German.

XX CC This invention describes a novel nucleic acid encoding a plant nucleobase
CC transporter (I). (I) is produced by complementation of a nucleobase
CC transport (NBT)-defective host cell with a plant gene bank by selection
CC of NBT-positive cells. (I) is used to isolate homologous sequences from
CC bacteria, fungi, plants, animals and humans, for expression of the
CC encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting

CC expression of (II) (when in antisense orientation), and to produce
CC transgenic crop plants. The transgenic plants have modified nucleobase
CC transport properties, e.g. altered affinity and substrate specificity
CC that may result in more efficient nucleobase transport in leaves, changes
CC in apical dominance, flowering behaviour and senescence, or improved
CC distribution of pesticides. This sequence encodes the Arabidopsis
CC thaliana PUP1 protein which is described in the method of the invention

XX SQ Sequence 1145 BP; 309 A; 234 C; 226 G; 376 T; 0 U; 0 Other;

Query Match 10.9%; Score 116.6; DB 3; Length 1145;
Best Local Similarity 48.3%; Pred. No. 4.2e-23;
Matches 434; Conservative 0; Mismatches 444; Indels 21; Gaps 3;

Qy 82 GTCTCTACTTCAACAATGGCGGTAAAGGATTTGGTCTCTACGTTTCTTTGAAACTGCAG 141
Db 140 GACTTTTACTAGAAAAAGCGGTAAAGACATGGCTCGAAACCTTGGTTTTCAGCTTGTAG 199
Qy 142 GCTTTTCCTGTATCTTCATTCCTCTGCTCTTCTTACATTACCCGCGGAGAAAGCAACA 201
Db 200 GGTTCCTCTTAAACCTTCCCTGTATATTAATTAAGCCTGAGCCGTCAGACTAAAA 259
Qy 202 ATGTGGGTGATAGTACAAAGTTTCTTTCTTATCAAAACCGCGTCTTTTATCGCCGCTGTA 261
Db 260 CCATTACCAAAAAAATACTTCTTCTTCTTGACA-----CTATCTTTAGTGATATTG 313
Qy 262 TTGTAGGCATTTCTCAGGGTTTGTATTAACACTACTTGTATGATGATGATAGCTTATCTTC 321
Db 314 GACTTGGCTTGTGTTGGCTGGACATTTGTAATTTGTACTCATTTGGGCTACTTTTACCTTC 373
Qy 322 CAGTTTCTACAGCTGCTCTTATCATTCGTTTACAGTTTGTAGTTTTTATAGCTATCTTTCAT 381
Db 374 CTGTCTCAACTTCTCTTTGATCTCTGCTGCGCAATTTGGCTTTTAAACGCGCTTCTTCTT 433
Qy 382 TCTTCATGGTTTAAACATAAAAGTTTCACTCTTTTACCATCAATGCTGTGTGTGTGACTG 441
Db 434 ACTTCTCTAAACTCACAATAAAATCACACCAATTTACTCAATTCACATTTGTTCTCTTAAACCA 493
Qy 442 TTGGTCTGCGGTTTTGGGAATGCATACCGAAACTGA-----TAAGCCAGTTC 489
Db 494 TATCTTCTACACTTCTTGTATCCACATGAACAGCAATCTCCCTCTTCTTACTTCAAGT 553
Qy 490 ATGAGACTCAACAGCAGTACATAAATGCTTTCTTGTATTAATCTGTAGCAGCAGCTTATGT 549
Db 554 CCGCAGCAAGTCCAAGTATGTGATGATACATCTCGCGGCTCGGTAGCTCAGCTGGTT 613
Qy 550 ATGCTTTTCATCTTGCCATTAGTGGAACTTGTCTTACCAGAAAGCTTAAGCAAAACCATGAGCT 609
Db 614 ATCTCTGGTGTCTTCTTTAACAGATTAACGCTTTCGAAAGATTTCTAAAGAAATACACAT 673
Qy 610 ATACCTTGTCTCGAGTTCAGTTTCCAGTTGATTTTGTGCTCTCTTGTCTTATTTGTCAGGCTCA 669
Db 674 TCAAGGCTATTTTAGACATGCCCCACATATCCGCTCTATGGTAGCTACTTGTGTAGTTGG 733
Qy 670 TCGGTATGTTTCATCGCTGGTGTATTTCAAGCAGGCTTTACCAAAAGAAAGCAAGAGTTCA 729
Db 734 TAGGACTTTTTGGAGTGGTGGTGGAAAG---CTGAGTACAGAAATGGAAGATTC 790
Qy 730 AGCTTGGAGAGGCAATTTCTATGTGTGGCTGTGTTTTTTCAGCCATCATATGGCAAGGCT 789
Db 791 AACTTAGGAAAAAGCTCATACATTTTGATAAACATCGGTTCAACGATATCATATGGCAAGCTT 850
Qy 790 TCTTCTTGGGAGCCATTTGGATTAATCTTCTCCACATCGTCTCTCGTCTCGGTTATTATGA 849
Db 851 GTTTGATTTGGAAGTGTGGTTTGAATATCGAAGTTTCATCGCTTTTTCCAATGTCATAA 910
Qy 850 TATCAGTGTCTTTTCCCAATTTACAGAGTTTATAGCTTTATATCTTACCATGAAAGTTTC 909
Db 911 GCATCTTTTGTACCAAGTTGTGCTGTTCTTGTCTTCTTCTTCCGTGATGAGATGA 970
Qy 910 AAGCTGAGAAAGGCACTTTCTTGTCTCTCTCCCTTTGGGCTTTGTCTCTTACTTTTAT 968
Db 971 GTGGAATCAAGTTGGTTGCAATGTTTTTGGCCATCTGGGATTTGTTTCTTATGGTTAT 1029

RESULT 15
AAC51119
ID AAC51119 standard; DNA; 1292 BP.
XX
AC AAC51119;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 67344.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 29-JUN-1999; 99US-0141287P.
PR 30-JUN-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 01-JUL-1999; 99US-0142055P.
PR 02-JUL-1999; 99US-0142390P.
PR 06-JUL-1999; 99US-0142803P.
PR 08-JUL-1999; 99US-0142920P.
PR 09-JUL-1999; 99US-0142977P.
PR 12-JUL-1999; 99US-0143542P.
PR 13-JUL-1999; 99US-0143624P.
PR 14-JUL-1999; 99US-0144005P.
PR 15-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144633P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148664P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.

PR	20-AUG-1999;	99US-0149722P.	
PR	20-AUG-1999;	99US-0149723P.	
PR	20-AUG-1999;	99US-0149929P.	
PR	23-AUG-1999;	99US-0149902P.	
PR	23-AUG-1999;	99US-0149930P.	
PR	25-AUG-1999;	99US-0150566P.	
PR	26-AUG-1999;	99US-0150884P.	
PR	27-AUG-1999;	99US-0151065P.	
PR	27-AUG-1999;	99US-0151066P.	
PR	27-AUG-1999;	99US-0151080P.	
PR	30-AUG-1999;	99US-0151303P.	
PR	31-AUG-1999;	99US-0151438P.	
PR	01-SEP-1999;	99US-0151930P.	
PR	07-SEP-1999;	99US-0152363P.	
PR	10-SEP-1999;	99US-0153070P.	
PR	13-SEP-1999;	99US-0153758P.	
PR	15-SEP-1999;	99US-0154018P.	
PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154799P.	
PR	22-SEP-1999;	99US-0155139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	29-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157753P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159283P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	

Query Match	10.9%;	Score 116.6;	DB 3;	Length 1292;
Best Local Similarity	48.3%;	Pred. No. 4.4e-23;		
Matches 434;	Conservative	0;	Mismatches 444;	Indels 21;
			Gaps	3;

Qy	82	GTCTCTACTTCAACAATGGCGGTAAAGGATTTGGTCTCTACGTTTCTTGAAGTCTGAG	141
Db	191	GACITTTACTAGCAAAAGCGGTAAAGACACATGGCTCGAAACCTTGGTTTCTGAGTCTGAG	250
Qy	142	GCTTTCTCGTTTATCTTCATTTCTCTGCTCTTCTTACATTAACCGCGCGGAGAGCAACA	201
Db	251	GGTTTCTCTTTAACCCCTTCCCTTGGCTATTATTAATTAAGCCCTGAGCCGTCGAAGACTAAA	310
Qy	202	ATGTGGGTGATGATGACAAAGTTTCTTTCTTATCAAAACCGCGTCTTCTTATCGCCGCTGTTA	261

Search completed: November 1, 2004, 07:20:11
Job time : 566.296 secs

Db	311	CCATTACCAAAAAAATACTTCTTCTCTTGACA-----CTATCTTTAGTGATATTG	364
Qy	262	TTGTAGGCATCTCTCAGGGTTTGATAAATACTTTGTAATGATATGGTATAGCTTATCTTC	321
Db	365	GACTTTGGCTTGTGTGTGGACATTTGATTTTGTACTCATTTGGGCTACTTTACCTTC	424
Qy	322	CAGTTTCTACAGCTGCTCTTATCATTTGCTTTCACAGTTTGTAGCTTTTATAGCTATCTTCTCAT	381
Db	425	CTGTCTCAACTTCTCTCTTTGATCTCTGCGTCGCAATTTGGCTTTTAAACCGCTCTTCTCTT	484
Qy	382	TCITTCATGGTTAAACATAAAGTTCACTCCTTTTACCATAATGCTTGTGTGTGTGACTG	441
Db	485	ACTTCTTAATACTCACAAAAAATCACACCAATTTATACTCAATTCACTTGTTCTCTTAACCA	544
Qy	442	TTGGTGTCTGGGTTTTTGGGAATGCATACCGAAACTGA-----TAAGCCAGTTC	489
Db	545	TATCTTCTACACTTCTTGTGTTATCCAAATGCAATGCAACAGCAATCTCCCTCTTCTACTTCAAGT	604
Qy	490	ATGAGACTCAACAGCAGTACATAACTGGTTTCTTGATTACTGTAGTAGCAGCAGCTGTTATGT	549
Db	605	CCGAGCCAAAGTCCAAATGATGATTTGGATACATCTCGCGGTGGTGGTGTAGCTCAGCTGGTT	664
Qy	550	ATGCTTTTCATCTTGCCATTAGTGGAACTTGCTTACCAGAAAGCTTAAGCAAAACCATGAGCT	609
Db	665	ATTCTCTGGTGTCTTCTTTAAACAGATTACGGGTTTCGAAAGAGATTCTAAAGAAATACACAT	724
Qy	610	ATACCTTGTGCTCGAGTCCAGTTGATTTTGTGTCTCTCTTGTCTTATTTGTCTAGCGTCA	669
Db	725	TCAAGGCTATTTTAGACATGGCCACATATCCGTCTATGGTAGCTACTTGTGTAGTTGTGG	784
Qy	670	TCGSTATGTTTCATCGCTGGTGATTTCAGAGCAGGCTTTACCAAAAGAACGAGAGAGTTCA	729
Db	785	TAGACTTTTTTGGAGTGGTGGTGGGAAAAG---CTGAGTACAGAAATGGAAGAGTTTC	841
Qy	730	AGCTTGGAGAGGCATTGTTCTATGTGTGGCTGTGTGTTTTTCAGCCATCATATGGCAAGGCT	789
Db	842	AACTAGGGAAGCTCATACATTTTGATAAAACATCGGTTCAACGATATCATGGCAAGCTT	901
Qy	790	TCITTCITGGAGGCCATTGGATTAATCTTCTCCACATGCTCTCTGCTCTCGGTTATTAAGA	849
Db	902	GTTTGATTGGAAGTGTGGTTTGATTATCGAAGTTTCATCGCTTTTTTCCAAATGTCATAA	961
Qy	850	TATCAGTGTCTTTTGCCCAATTACAGAGTTTTTAGCTGTATATTTCTACCATGAAAAGTTTC	909
Db	962	GCACTCTTGTTTTACCAAGTTGTGCGCTGTCTTGTGTGTCTTCTTCCGTAATGAGAAGA	1021
Qy	910	AAGCTGAGAAAGGACTTTCTCTTGTCTCTCTCCCTTTGGGGCTTTGTCTCTTACTTTTTAT	968
Db	1022	GTGGAATCAAGTTGGTTTGCAATGTTTTTGGCCATCTGGGGAATTTGTTTCTTATGCTTAT	1080

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:45:29 ; Search time 3851.17 Seconds
(without alignments)
10228.408 Million cell updates/sec

Title: US-09-913-767-6
Perfect score: 1081
Sequence: 1 caaatccaacagtccaagat.....ataagtgatatttgaaag 1081

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	858.8	79.4	1519	CNS0A3B5	BX826795 Arabidops
2	433.6	40.1	822	BZ491815	BZ491815 BONPU42TF
3	425.4	39.4	926	BX835351	BX835351 BX835351
4	424.2	39.2	849	BX838927	BX838927 BX838927
5	369.6	34.2	1340	CNS0A29J	BX827536 Arabidops
6	369.2	34.2	668	BQ704880	BQ704880 Bn01_01n0
7	366.6	33.9	1375	CNS0A2S8	BX827187 Arabidops
8	363.4	33.6	1431	CNS0A2XR	BX829081 Arabidops
9	357.2	33.0	1312	CNS0A2U8	BX827470 Arabidops
10	331.6	30.7	1359	CNS0A2R3	BX827477 Arabidops
11	299.8	27.7	768	AJ795195	AJ795195 AJ795195
12	284.2	26.3	1323	CNS0A3GP	BX826811 Arabidops
13	279.4	25.8	666	BQ157452	BQ157452 NF105A121
14	269.8	25.0	922	BZ967767	BZ967767 PUDG56TD
15	263	24.3	782	CO105310	CO105310 GR_Bb003
16	259	24.0	480	AV522367	AV522367 AV522367
17	256.4	23.7	627	BQ155273	BQ155273 NF078F121
18	255.6	23.6	770	CG344371	CG344371 OG4AG15TC
19	253.4	23.4	1374	CNS0ABEV	BX814955 Arabidops
20	249.6	23.1	837	CB292743	CB292743 UCRCS01_0
21	247.8	22.9	805	BX927581	BX927581 BX927581
22	247.4	22.9	699	B77316	B77316 T3213TF TAM
23	242.4	22.4	691	BZ087060	BZ087060 lkh10d12.
24	240	22.2	399	AV559386	AV559386 AV559386

25	233.2	21.6	630	4	BG525645	BG525645 52-8 Stev
26	233	21.6	739	7	CO105307	CO105307 GR_Bb003
27	221.4	20.5	1289	3	CNS0A2HI	BX837065 Arabidops
28	220.4	20.4	765	1	AJ795240	AJ795240 AJ795240
29	218.6	20.2	762	5	BQ121182	BQ121182 EST606758
30	217.2	20.1	800	6	CF069001	CF069001 EST669722
31	214.6	19.9	568	4	BMI76948	BMI76948 saJ74H01.
32	213.4	19.7	562	7	CK759270	CK759270 pam01-13m
33	206.4	19.1	809	7	CF445584	CF445584 EST681929
34	205.4	19.0	1299	3	CNS0A41J	BX826394 Arabidops
35	203.4	18.8	735	4	BM408950	BM408950 EST583277
36	203	18.8	509	1	AI729914	AI729914 BNLGH1568
37	202.2	18.7	684	1	AJ800150	AJ800150 AJ800150
38	201.8	18.7	806	4	BI310188	BI310188 EST531193
39	200.2	18.5	762	4	BM413141	BM413141 EST587468
40	193.4	17.9	652	6	CA255028	CA255028 SCEPFL417
41	192.6	17.8	589	7	CK096760	CK096760 UB25CPF10
42	191.2	17.7	765	7	CNI46235	CNI46235 WOUN0139
43	190.2	17.6	676	6	CA920055	CA920055 EST637773
44	189.6	17.5	761	7	CNI25253	CNI25253 RHOH1_10
45	189.2	17.5	633	1	AI777352	AI777352 EST263760

ALIGNMENTS

RESULT 1
LOCUS CNS0A3B5 1519 bp mRNA linear HTC 06-FEB-2004
DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFB632D12 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).
ACCESSION BX826795
VERSION BX826795.1 GI:42460970
KEYWORDS HTC; GSLT_cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1519)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1519)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_projet_BF/Full length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

Location/Qualifiers
1. 1519
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"

[illegible]

Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation

JOURNAL COMMENT
Unpublished (2004)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: secref@genoscope.cns.fr, Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences).
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES source
Location/Qualifiers
1..849
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="GSLTFB632F12"
/tissue_type="Flowers and buds"
/clone_lib="Arabidopsis thaliana Flowers and buds Col-0"

ORIGIN
Query Match 39.2%; Score 424.2; DB 5; Length 849;
Best Local Similarity 80.6%; Pred. NO. 5.9e-104;
Matches 623; Conservative 0; Mismatches 78; Indels 72; Gaps 8;

Qy 1 CAATCCACAGTTCACAGATGAAGAAATTCAGTCAGTACGAGCCCAAGCAAGTATCTC 60
Db |||||||
Qy 86 CAATCCACAGTTCACAGATGAAGAAATTCAGTCAGTACGAGCCCAAGCAAGTATCTC 145
Db |||||||
Qy 61 ACTCTAACACATACAAACGGTGGCTCAG-----GAGTATCTATGACAGCGGAGGAGGAGTAA 88
Db |||||||
Qy 146 ACTCTAACACATACAAACGGTGGCTCAGCGGTGACTTTATATACATCTTTGTCATTTCAG 205
Db |||||||
Qy 89 -----GAGTATCTATGACAGCGGAGGAGGAGTAA 120
Db |||||||
Qy 206 GCCAAACAGTTGCTACAAATTTGGCGCAGAGTATCTATGACAGCGGAGGAGTAA 265
Qy 121 TGGCTAGCAACGGTAGTTCAACTGTTGGCTTTCTGCTGCTACTTCCATATATATCTTG 180
Db |||||||
Qy 266 TGGCTAGCAACGGTAGTTCAACTGTTGGCTTTCTGCTGCTACTTCCATATATATCTTG 325
Db |||||||
Qy 181 TCATTTAAACACATGCAACCACTGATAGATGGAAGAAAGAACTCACCCTA--GGAACC 238
Db |||||||
Qy 326 TCATTTAAACACATGCAACCACTGATAGATGGAAGAAAGAACTCACCCTAGTGAATCC 385
Qy 239 GTGATATGGTTTACGTAGTGTGAC--TTCTGTAGGAGCAGATGCTATCTGTACTCC 297
Db |||||||
Qy 386 GTGATATGGTTTACGTAGTGTGAC--TTCTGTAGGAGCAGATGTTGATCTGTACTCC 445
Qy 298 ATTGACATCTTTACTTACCGGTTTCTACCTATTCCTGATCTGTGATCTCAGTTAGCC 357
Db |||||||
Qy 446 TTGTGCTCTCTTACTTTCCCTTTTCTCC---TTTCTGATCTGCTCTCTGTGATCC 502
Qy 358 TTCAATGCTCTTCTCTTATTTTCTTAACACAAAAAAGTACCCTCATCATTTTAAAT 417
Db |||||||
Qy 503 TTTRA--GTGCTACGCTTTGTGGTAACTCAC-AAAACTTACCCCTCATCTTTAAA- 558
Qy 418 TCTCTTTTCTCTTAACTATATCTTCCACCCTACTTGCAATCAATATGAGGAGACAGAC 477
Db |||||||
Qy 559 TCTCTTTTCCCTTAACGATTA--CTCTCTCCGAGCTTGTATTCAATTAATGAGGAGACATTC 616
Qy 478 TCCACAAAGTTTACAAAGGAGATGTCTCAAAGGTTTCATATGACCGTTGCTGCGTCT 537

Db |||||
617 TCCATATAGATATAAAGATATTATTTTAAGTTACAGATGCACCGTGTGCGCCC 676
Qy 538 GCTGGTTATGCTAGTCTTTATCCCTACAACAGCTAGCCTTTCTAAAAGTCTTAAAGAG 597
Db |||||
677 GCGGTAATGCTCGAGTCTTATCCGACAATATCCAGCCTTTCTAAAAGTCTTAAAGAG 736
Qy 598 CAAATTTCTCAGAGTTATGGATATGATATATCTACGTGAGTCTAGTGCCAGTTGTGTT 657
Db |||||
737 CATATTTCTCATATGATGGATTTGATAATCTTTGTGATACACAGTGGCTGTAGAGTT 796
Qy 658 AGCGTGTGGGGCTTTTGTAGCAGTGAAGTGGAAAACTTTGAGCACTGAAAT 710
Db |||||
797 AGCGTGAAGGGCTTTAGTCAGTCAGTGAAGTGGAACTTTGAGTAGTGTAAAT 849

RESULT 5
CNS0A29J 1340 bp mRNA linear HTC 06-FEB-2004
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLT5712A05 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).

ACCESSION
VERSION BX827536.1 GI:42459416
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation

Unpublished
2 (bases 1 to 1340)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr

COMMENT
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=arabidopsis>.

FEATURES source
Location/Qualifiers
1..1340
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
1..1340
/gene="At4g18197"

ORIGIN
Query Match 34.2%; Score 369.6; DB 3; Length 1340;
Best Local Similarity 62.2%; Pred. No. 4.6e-89;
Matches 601; Conservative 0; Mismatches 359; Indels 6; Gaps 1;

Qy	89	GAGTATACTATGACNACGGAGAAACAGTAATGGCTAGCAACCGTAGTTTCAACTGTTG	148
Db	225	GATTGTACTATGAAATGGAGAAATAGCACAATATGTGGTAACACTCTTCTTCAACTCAATTG	284
Qy	149	GCTTTCCTGTGCTACTCTCCATATTATATCTTGTCAATTTA-----AAACACATGCAACAA	202
Db	285	GCTTCCCTGTACTGGTTCTGTTCGGCTCTTTCTCGAATCAGGCNACCCAAATCAACAG	344
Qy	203	CTGATAGAGATGGAAAAAGAACCTCACTCAGAGAACCGTGTAATGGTTTACGTAGTGCTG	262
Db	345	ATACAAAATGTCAGTCAGTCGCCCTTCTTCCACACCCCTTGCATCGGTTTACTTGTGCACTG	404
Qy	263	GACTTCTGTAGGAGCAGATGCTATCTGTACTCCATTGGACTTCTTTTACTTACCCGTTT	322
Db	405	GACTGCTAGTGTCCGCTTATGCTTATTTGTCTGCAGTAGGGTGTCTACTTTACAGTCT	464
Qy	323	CTACCTATTCCTGATCTGTGCACTCTCAGTTAGCTTCAATGCTTTCTTCTTATTATTTTC	384
Db	465	CTACTTTCTCCTCATCTTGGCCCTCAGAGTTGGCCTTCACTGCGCTTTTCTCATATTTCC	524
Qy	383	TTAACTCACAAAACTTACCCCTATCATTTTAAATTCCTTTTCTCCTTAACTATATCTT	442
Db	525	TTAACTCGCAAAAGTTCACTCCTTTGATGAGTCAGTTCTTTCGCTTCTCCTCACTGATCCT	584
Qy	443	CCACCTACTTGCATTTCAATTAATGAGGAGACAGACTCCACAAAAAGTTTACAAAAAGGAGT	502
Db	585	CTGCTCTTTTGGTCAACACTGATTCAGAAAACTCAACTAATGTATCTAGAGTACAGT	644
Qy	503	ATGTCAAAGGTTTCATATGCAACCGTTGCTGGCTCTGCTGGTTATGGTCTAGTCTTATCCC	562
Db	645	ATGTCATCGGGTTTCATATGATACCACTCGGTGCTCCGCTGGGATTTGGACTGTTACTATCTC	704
Qy	563	TACAAACAGCTAGCCCTTCTTAAAGTCTTAAAGAACCAAAATTTCTCAGAACTTATGATA	622
Db	705	TGATACAAATGCTCTTTCAGGAAAGTTTTCAGAAAGCATACATCTCAGCAGTCAGGACT	764
Qy	623	TGATAATCTACGTGAGTCTAGTGCCAGTTGTGTGTAGCGTGGTGGGCTTTTTCGTAGCA	682
Db	765	TGGCATTATACAGTCTCTAGTTGCGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAAGTG	824
Qy	683	GTGAGTGGAAACCTTTCAGGACAGTCAAAATGGATAACTACAAACATGGGAAGTATCTTACA	742
Db	825	GAGAGTGGGAAACCTTTCGCAAGTCAGATGAGAAACTACAACTCGGAAAGTGTCTATGTG	884
Qy	743	TTATGAACCTAGTGTGGACAGCTGTTACCTGGCAGTTATTCCTCATCGGTGGCACAGGAC	802
Db	885	TTTTGACTTTAGCCTCGGCAGCTATTTCTCGCAAGTCTACACTCTTGGTCTTGTGGGAT	944
Qy	803	TGATCTTCGAGCTCTCCTCTCTATCTTCAAAATGCAATAAGCGTTTTCGGACTCCCACTGG	862
Db	945	TGATCTTCGAGTCATCTCTGTGTCTTCTTCCAAATCCATAACAGCTGTGGGATTTGCTATAG	1004
Qy	863	TTCTATCTTCGGCTGTAAATCATTTTTCATGACAAAAATGAATGGCTTAAAGGTGATTTCTA	922
Db	1005	TTCCAGTTTCGGCAGTGATAGTTTTCATGATAGATAGGACGATCAAAATCTTCTCCA	1064
Qy	923	TGATCTAGCTATTTTGGGGTTTCACTTCTATGTCACCAACAATATCTTGAAGCAAAA	982
Db	1065	TTATTTTAGCTATCTCGGGCTTCTCTCATTTCTGTCATCTAGCACTACCTCGACGAAAAAGA	1124
Qy	983	ACTTGAAGAAAAATCATGAAATCAACAACAGAAATCCCCTGACCCACCAAGACGAGAAG	1042
Db	1125	AGTTGATATCTAGCCACACAAGTCTGTAGGAGATCTTCACTACCTGTTGGAGAAAGTC	1184
Qy	1043	AGTCAA	1048
Db	1185	ACAAA	1190

RESULT 6
BQ704880
LOCUS

DEFINITION Bn01_0ln06_A
 Bn01_AAPC_ECORC transgenic Brassica napus overexpressing BNCBF17, co-
 nstitutively frost_tolerant Brassica napus cDNA clone Bn01_0ln06,
 mRNA sequence.

ACCESSION B0704880
VERSION BQ704880.1 GI:21844299
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 668)
AUTHORS Singh, J., Allard, G., Tinker, N., Robert, L., Lacroix, C., De Moors, A.,
 Chagnon, J., Farah, S., Couroux, P. and Hattori, J.
TITLE Expressed Sequence Tags from constitutively frost tolerant
 transgenic Brassica napus overexpressing BNCBF17
JOURNAL Unpublished (2002)
COMMENT Contact: Singh, J.A.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-food Canada
 KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
 0C6, Canada
 Tel: (613) 759-1662
 Fax: (613) 759-1701
 Email: singhja@agr.gc.ca.
 Location/Qualifiers
 1..668
 /organism="Brassica napus"
 /mol_type="mRNA"
 /cultivar="Westar"
 /db_xref="taxon:3708"
 /clone="Bn01_0ln06"
 /tissue_type="fourth leaf"
 /dev_stage="3 weeks seedling grown at room temperature"
 /clone_lib="Bn01_AAPC_ECORC transgenic Brassica napus over
 expressing BNCBF17, constitutively frost tolerant"
 /notes="vector: Bluescript SK-/XhoI-EcoRI; Site 1: EcoRI;
 Site 2: XhoI; Germinated in soil flats and seedlings grown
 for 3 weeks in a Conviron E-15 cabinet set at 20°C /16 hr
 light (250 Em-2sec-1) and 16 °C/ 8 hr dark. Fourth leaves
 collected at 9 am and immediately frozen."

ORIGIN
 Query Match 34.2%; Score 369.2; DB 5; Length 668;
 Best Local Similarity 79.0%; Pred. No. 5e-89;
 Matches 433; Conservative 22; Mismatches 90; Indels 3; Gaps 2;

Qy 524 CCGTTGCTCGGTCTGCTGGTTATGGTCTAGTCTTATCCCTACACAGTAGCTTCTTAA 583
 Db 1 CCATTGGTGCTTCTGCTGGGTATGGCTTAAGCTTATCTCTVACAAAGTAGTCAGTA 60

Qy 584 AAGTCCTAAGAGCAAAATTTCTCAGAAGTTATGATATGATATTAATCTAGCTAGCTAG 643
 Db 61 AAGTCCTTAAAGGCGCAAACTTCTCSAAGTCATGGAATTSATCATCTACGTAAGTCTAG 120

Qy 644 TGGCCAGTCTGTTAGCTGGTGGGCGCTTTTCTAGCAGTAGTGCGAAACCTTCGAGCA 703
 Db 121 TGGCGWCTGTGTAAAGGTGGTGGGCTCTTTGCTAGTGGAGAGTGGGAAACTTTSWGA 180

Qy 704 GTGAATGGGATAACTACAAACATGGGAAGTATCTTACATATTATGAACCTAGTGTGACAG 763
 Db 181 GTGAATGGATAGCTACAAGCTTTGGAAGGTAATCTTACGTTATGAACCTAGTGTGACGG 240

Qy 764 CTGTTACCTGGCAGTTATCTTCATCGGTGGCAGCAGACTGATCTTTCGAGCTCTCT-CT 822
 Db 241 CTGTTACCTGGCAGGTATTCAACATTTGGTAGCAGTGGGCTTATCTTCGAGATTTCTCT 300

Qy 823 CTATTTCTCAATGCAATTAAG- -CGTTTGGGACTCCCGAGTGGTTCCTATCTTGGCTGTAA 880
 Db 301 CTATTTCTCAANYGCTATGTGGCGTGTTGGGTTGGCTCTGGGCTCCGCACTGGCTGTC 360

Qy 881 TCATTTTCCATGACAAAATGAATGGGTAAAGGTGATTTCTATGATCTCCTAGCTATTGGG 940

```

||||| 361 TCATTTTCATGATAAATGATGSGTCMAGSTGATATCTATGATCTCYSGCTMYCTGGG 420
||||| 941 GTTTCACCTTCATGCTACCAACAATATCTTGATGACAAAACCTTGAGAAAATCATG 1000
||||| 421 STYTCATGTCCTATGATATACCAACACTATCTTGATGACAAAACCTTGAGAAAAGCTTG 480
||||| 1001 AAATCACACAACAGAAATCCCTGACCCACCAAGACGAGAGTCAACTTGGCNAATCAA 1060
||||| 481 GAATCCCAACAACCGGAYCTCCGACTCACCAAGAACGAAAAGGGYCMAGTCGCAAAAAGA 540
||||| 1061 AATAAGCT 1068
||||| 541 TACAACT 548

```

RESULT 7

CNS0A2S8 1375 bp mRNA linear HTC 06-FEB-2004
 Arabidopsis thaliana Full-length cDNA Complete sequence from Clone
 GSI7LS20ZA10 of Adult vegetative tissue of strain col-0 of
 Arabidopsis thaliana (thale cress).

ACCESSION

VERSION BX827187.1 GI:42459879

KEYWORDS

HTC; GSI7 cDNA.

SOURCE

ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
 Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
 Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
 Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 1375)

AUTHORS

Genoscope.

TITLE

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

COMMENT

The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen)
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES

Location/Qualifiers

```

1..1375
/mol_type="mRNA"
/organism="Arabidopsis thaliana"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSI7LS20ZA10"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
1..1375
/gene="At4g18205"

```

gene

1..1375

ORIGIN

Query Match 33.9%; Score 366.6; DB 3; Length 1375;
 Best Local Similarity 62.7%; Pred. No. 3e-88;
 Matches 599; Conservative 0; Mismatches 344; Indels 6; Gaps 1;

```

QY 59 TCACTCTAACACATACAAAACGGTGGCTCAGGAGTATACTATGACAAACGGAGGAACAGTA 118
DB 193 TCTGCCAACCACTTGGCTACAGTCTCGGGTAGACTGTACTATGAAATGGAGGGAAAAGCA 252
QY 119 AATGGCTAGCAACGGTAGTTCAACTCTGTGGCTTTCCTGTGCTACTTCCATATATATATCT 178
DB 253 CATATGTGGTAACACTTCTTCACTCAATGGCTTCCCTGTACTGATTTCTCTCCGCTTCT 312
QY 179 TGTCAITTTA-----AAACACATGCAACAACTGATGAGATGGAAAAGAACCTCACCTA 232
DB 313 TTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATGTCAAGTCAGTCCTCCCTTCTTCA 372
QY 233 GGAACCGGTGATTTGGTTTACGTAGTGTGGACTTCTGTAGGAGCAGATGCTATCTGT 292
DB 373 CCACCTTGCATCGGTTTACTTGTGCACGTGCTAGTGTCTGCTTATGCTTATTTGT 432
QY 293 ACTCATTTGGAATCTTTTACTTACCGTTTCTACTATTCCTCGATCTGTGCACTCTCAGT 352
DB 433 CTGCAGTTGGTGTGCTTTACTTACCAGTCTCTACTTTCTCCCTCATCTTGGCCTCACAGT 492
QY 353 TAGCCTTCAATGCTTTCTCTTATTTCTTAATCTCACAATACTTACCCCTATCATTTT 412
DB 493 TGGCCTTCACTGCTCTTTTCTCATATTTCTTAATCTGCAAAAGTTCACTCTTTGTATG 552
QY 413 TAAATTTCTTTTCTCTTAACTATATCTTCCACCTACTTGCATTGCAATAATAGGAGA 472
DB 553 TCAATTTCTTTGTTTCTCTTACGGTTTCTCTGCGCTTCTCGTGGTCAACACTGATTCAG 612
QY 473 CAGACTCCAAAAAGTTTACAAAAGCAGATGTGCAAAAGTTTTCATATGACACCGTTGCTG 532
DB 613 AAAACACACAAATGATCTAGAGTACAATATGTGATTTGGTTTCATCTGTACCATTTGGT 672
QY 533 CGTCTGCTGTTATGCTCTTATCTTATCCCTACAAACAGCTAGCCCTTTCTTAAAGTCTTAA 592
DB 673 CTTCGGCTGGGATGGATTTGGTACTATCTCTGATACAACTGCTCTTCAGGAAAAGTTTTC 732
QY 593 ARAAGCAAAATTTCTCAGAAATGATGATATGATTAATCTAGTGAAGTCTAGTGCCAGTT 652
DB 733 CGAAGCATACATCTCTCAGCAGTCTCTGCTTGTCCAATACCAGTCTCTAGTTGCAACTT 792
QY 653 GTGTTAGCTGGTGGGCTTTTCTGCTAGCAGTGAAGTGAAGAACTTTGAGCAGTGAATGG 712
DB 793 GTGTGTTACTCATAGGACTGTTTCCAAGTGGAGATGGAGAACTCTGCCAAGTCAGATGA 852
QY 713 ATAACTCAAAACATGGGAAGGTATCTCATATTATGAACCTAGTGTGGACAGCTGTTACCT 772
DB 853 GAAACTACAAACTGGGAAAAGTGTATATATCTTTGACTTTGGCCTCAGCAGCTATTTTCT 912
QY 773 GGCAATTTATCTCCATCGGTGGCACAGACTGATCTTCGAGCTCTCCTCTTATTTCTCAA 832
DB 913 GGCAAGTATACACTGTTGGTGTGGGATTAATCTTCGAGTCTTCTCTGTGTCTCCA 972
QY 833 ATGCAATAAGCGTTTGGGACTCCAGTGGTTCCTATCTTGGCTGTAATCATTTTCCATG 892
DB 973 ATTCCATAACTGCTGTGGACTGCTTATGTTCCAGTTTGTAGCAGTGTATGTTTCCATG 1032
QY 893 ACAAATGAATGGGTAAAGGTGATTTCTATGATCTCTAGCTATTTGGGGTTTCACTTCT 952
DB 1033 ATAAGATGGATGCATCAAGAAGATTTTCTCCATCATTTTAGCTATCTGGGGCTTCTCTCAT 1092
QY 953 ATGTCTACCAACAATATCTTGATGACAAAACCTTGAAGA 991
DB 1093 TTGTTCTATCAGCACTACTCTGACGAAAAGAGTTGAAGA 1131

```

RESULT 8

CNS0A2KR

LOCUS

DEFINITION

Arabidopsis thaliana Full-length cDNA Complete sequence from clone

GSI7LS1692B03 of Silique of strain col-0 of Arabidopsis thaliana

(thale cress).

ACCESSION

BX829081

CNS0A2KR 1431 bp mRNA linear HTC 06-FEB-2004
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 GSI7LS1692B03 of Silique of strain col-0 of Arabidopsis thaliana
 (thale cress).
 BX829081

full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
[http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length)

<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES

Location/Qualifiers

1..1312

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="Col-0"

/db_xref="taxon:3702"

/clone="GSLTSL52F07"

/tissue_type="Adult vegetative tissue"

/plasmid="pCWVSPORT_6"

1..1312

/gene="At4g18205"

gene

ORIGIN

Query Match 33.0%; Score 357.2; DB 3; Length 1312;
Best Local Similarity 62.8%; Pred. No. 1.1e-85;
Matches 590; Conservative 0; Mismatches 343; Indels 7; Gaps 2;

QY 59 TCACCTTAACATCAACACGGTGGCTCAGGAGTATACATGACAAACGGAGAAACAGTA 118
DB 207 TCTGCCAACCACTTGTACAGTTCCTGGGTAGACTGTACTATGAGAATGGAGGAAAGCA 266

QY 119 AATGCTAGCAGCGTAGTCAACTTGTGGCTTCTCGTGTACTTCCATATATATCT 178
DB 267 CATATGTGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGATTTCTTCCGCTTCT 326

QY 179 TGTCTATTTA-----AAACACATGCAACCACTGTAGAGATGGAAAGAAACCTCACCTA 232
DB 327 TTCTCGAATCAGGCACCAATCAACAGATACAAATTCAGTCAGTCCCTTCTTCA 386

QY 233 GGAACCGTGTATGTTTACGTAGTGTGGACTTCTTGTAGGACGATTCGTATCTGT 292
DB 387 CCACCGCTTGCATCGGTTTACTTGTGCACCTGGACTGTCTGCTTATGCTTATTTGT 446

QY 293 ACTCATTTGGACTTCTTTACTTACCGTTTCTACCTATTCCTGATCTGTCATCTCAGT 352
DB 447 CTGCAGTTGGGTGTCTTACTTACCAGTCTTACTTCTCCCTCAATCTTGGCCTCACAGT 506

QY 353 TAGCCTTCAATGCTTTCTTCTTCTTATTTCTTAACTCAAAACCTTACCCCTATCATTT 412
DB 507 TGGCCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAGTTCACTCTTTGATAG 566

QY 413 TAAATCTCTTTTCTTCTTAACTATATCTTCCACCTACTTGCATTCATTAATAGAGAGA 472
DB 567 TCAATTTCTTTTCTTCTTACGTTTCTCTGCTGCTCTCTGCTGCTCAACTGATTCAG 626

QY 473 CAGACTCCACAAAGTTACAAAGAGAGATGTCA-AGGTTTTCATATGACCGTTCT 531
DB 627 AAAACACAAATATATCTAGATGACAATATGTGATTTGGGGTTCATCTGACCAATGGT 686

QY 532 GCGTCTGCTGGTGTATGGTCTAGTCTTATCCCTTACAAACAGCTAGCCCTTCTTAAAGTCTCA 591
DB 687 GTTTCGCTGGGATTTGGATTGGTACTATCTGTATACAACTGCTCTTCAGGAAAGTTTTC 746

QY 592 AAGAAGCAAAATTTCTCAGAAAGTTATGATATGATAATCTACGTGAGTCTAGTGCCAGT 651
DB 747 ACGAAGCATACATCTCAGCAGTCTCTGAGTCTGGCAATTCACAGTCTCTAGTTGCAACT 806

QY 652 TGTGTTAGCGTGGTGGGCTTTTCTTCTAGCAGTGTGAGAACTTTTCAGCAGTGAATG 711
DB 807 TGTGTGGTACTCATAGGACTGTGTCAGATGGAGAGTGGAGAACTCTCCCAAGTGAGATG 866

QY 712 GATAACTACAAACATGGGAAGGTATCTTACATTATGAACTAGTGTGGACAGCTGTTACC 771

DB 867 AGAACTACAACTGGGAAAGTGTATATATCTTGAATTTGGCTCAGCAGCTATTTTC 926
QY 772 TGGCAGTATTTCTCCATCGGTGGCAGGACTGATCTTCGAGCTCTCTCTCTATCTCA 831
DB 927 TGGCAAGTATACACTGTGTGGTGTGTGGGATTAATCTTCGAGTCTTCTCTGTGTCTCC 986
QY 832 AATCAATAAGCGTTTGGGACTCCCACTGCTTCTATCTTGGCTGTAAATCATTTTCCAT 891
DB 987 AATTCATAACTGCTGTGGGACTCCCTATAGTTCAGTTGTAGCAGTATAGTATTTCCAT 1046
QY 892 GACAAATGAATGGGTTAAAGTGATTTCTATGATCTCTAGTATTTGGGGTTTCACTTCC 951
DB 1047 GATAGATGGATGATCAATCAAGATTTTCTCCATCATTTTAGCTATCTGGGCTTCTTTCA 1106

QY 952 TATGCTTACCAACAATATCTTGTATGACAAAACCTTGAAGA 991
DB 1107 TTTGTATCAGCACTACCTCGACGAAAGAAAGTTGAAGA 1146

RESULT 10
CNS0A2R3 1359 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTSL52E07 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX827477.1 GI:42459927
VERSION BX827477.1
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1359)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1359)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-Nov-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
[http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length)
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES Location/Qualifiers
1..1359
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL52E07"
/tissue_type="Adult vegetative tissue"
/plasmid="pCWVSPORT_6"
1..1359
/gene="At4g18197"

gene
ORIGIN


```
VERSION B0157452.1 GI:20294511
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eusteroids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 666)
AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL Medicago truncatula irradiated library
COMMENT Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 666 Std Error: 0.00
Plate: 105 row: A column: 12
Seq primer: TCACACAGGAAACAGCTATCAC.
FEATURES
    source
        1..666
            /organism="Medicago truncatula"
            /mol_type="mRNA"
            /db_xref="taxon:3880"
            /clone="NF105A12IR"
            /tissue_type="seedlings"
            /dev_stage="seedling"
            /clone_lib="Irradiated"
            /note="Vector: Lambda Zap; Seedlings were exposed either
to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
Gamma-irradiated samples were harvested at 6, 12, 24 and
48 hours after treatment. UV-irradiated samples were
harvested 24 hours post-treatment. cDNA was prepared from
polyA+ enriched, pooled samples of equivalent amounts of
total RNA from each sample. The cDNA was directionally
ligated into the Uni-Zap XR vector (Stratagene) and
packaged using the Gigapack III Gold packaging extracts.
Phagemids containing cDNA inserts were in vivo excised
from the recombinant Uni-Zap XR vector using ExAssist
helper phage and the E. coli strain XL1-Blue MRP'
(Stratagene). Excised plasmids were plated using SOLR
cells."
ORIGIN
Query Match 25.8%; Score 279.4; DB 5; Length 666;
Best Local Similarity 64.7%; Pred. No. 1.3e-64;
Matches 431; Conservative 0; Mismatches 232; Indels 3; Gaps 1;
QY 270 TGTAGAGCAGATGCTACTGTCTACTCCATGTCAGCTTCTTTACTTACCGGTTTCTACCTA 329
DB 1 TTTAGCAGGGGCAATTTGATGTATTTCCAAATGAGCTTCTATATCTTTCACTTTCCACCTA 60
QY 330 TTCCCTGATCTGTCATCTCAGTTCAGTTCGATCTTCAATGCTTCTTCTCTTATTTCTTAACCTC 389
DB 61 TTTCTTTACTATGTGCAACACAATTAGCCCTTCAATGAGATTTTCTTATTTCTTAAATTC 120
QY 390 ACAAAAACCTTACCCCTATCATTTTAAATTCCTTTTCCCTCTTAACTATATCTTCCACCCCT 449
DB 121 CCAAAAGTTTCACAGCATTATATTCATTTCTGTAGTTCTGTCTACTATTTTCTGCTGCCCT 180
QY 450 ACTTGCAATTCATATGAGAGACAC---AGACTCCACAAAAGTTACAAAAGGAGAGATGTT 506
DB 181 ACTTGCCATCAATGAGAGAGATTTCTGAGGACTCGACGAATCTTCCCTAAAGATAAGCGAGT 240
QY 507 CAAAGGTTTCATATGACACCGTTGCTGCGTCTGCTGTATGCTAGTCTTATCCCTACA 566
DB 241 ACTTGGATTTCGCTTCCACCCCTGGTGATCTGTATACATTTTCTTTATATATCTTTCCCTGTT 300
```

```
QY 567 ACAGCTAGCCTTTCTAAAGTCTCTAAAGAACGAAATTTCTCAGAAAGTTATGGATATGAT 626
DB 301 GCAGCTTTCTTTTGACAAAGTTATAAAAAGCGAAACCTTTTCTGTTATATTAGACATGCA 360
QY 627 ATCTACGTGAGTCTAGTGGCCAGTTGTTTACCGTGTGGGGCTTTTTCCTAGCAGCTGA 686
DB 361 ATTTTACCCCTCTTTTCGTCGCTACTTGTGCTTTGTGTTGTTGGTGTGTTTTCGAAGTGGAGA 420
QY 687 GTGGAAAACCTTTGAGCAGTGAATGAATACTTACAAACATGCGGAAGGTATCCTACATTAT 746
DB 421 ATGGAAGAAGTTTGAATAGGAGATGAAGAATATGACTGGGAAGGTCTCATATATAT 480
QY 747 GAACCTAGTGTGCGACAGCTGTTTACCTGGCAGTTTATTTCTCCATCGGTGGCAGGAGCTGAT 806
DB 481 GACTTTGTTTGGACTGCTGTGACATGCAAGTATCTTCCATTGGTATGTTAGGGGTAGT 540
QY 807 CTTGAGCTCTCTCTCTATTTCTCAATGCAATAGCGTTTGGGACTCCAGTGGTTCC 866
DB 541 TTTAGATGTATCTTCAATGTTCTCAATGTAATAGGTACAGTGGTACTTCCCGTANGTCC 600
QY 867 TATCTTGGCTGTAATCATTTTCCATGACAAATGAATGGTTAAAGGTGATTTCTATGAT 926
DB 601 TATCTTGTCTGTATTTTCTTCCATGACAAAATTAATGGGTGAAGTTTATAGCTTTGAT 660
QY 927 CCTAGC 932
DB 661 GTTAGC 666
RESULT 14
LOCUS BZ967767
DEFINITION BZ967767 922 bp DNA linear GSS 25-MAR-2003
PUDG56TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTa184115,
genomic survey sequence.
ACCESSION BZ967767
VERSION BZ967767.1 GI:29185853
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 922)
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Remnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
    source
        1..922
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBTa184115"
            /clone_lib="ZM 0.6 1.0 KB"
            /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
ORIGIN
Query Match 25.0%; Score 269.8; DB 8; Length 922;
Best Local Similarity 60.4%; Pred. No. 5.7e-62;
Matches 445; Conservative 0; Mismatches 292; Indels 0; Gaps 0;
QY 260 TTGGACTTCTTTGTAGGAGCAGATTGCTATCTGTACTCCATTGGACTTCTTTACTTACCGG 319
```

Db 3 TGGGCTCATATCTCTCGGATGATCTATGATCTATGGCTTTGTACCTTCTG 62
Qy 320 TTTCTACCTATTCCTCGATCTGTGCACTCAGTTAGCCTTCAATGCTTTCTCTTAT 379
Db 63 TATCAACATATTCGCTCTATTTGGCTAGCTAGCTGGCTTCAATGCTCTCTCATATG 122
Qy 380 TTCCTAACTCAGAAATCTACCCCTATCATTTTAAATTCCTTTCTCTTAATATAT 439
Db 123 TCCTAAATGCTCAAAAGTTCAACCCATCATTTTCACTCAGTAATCTCTTACCTTTT 182
Qy 440 CTTCCACCTCTCTGATTCATTAATAGGAGACAGCTCCACAAAAGTTACAAAAGGAG 499
Db 183 CGCTGCGCTCTTGAGTTGACGAGATTTCTCAGGTACCAATGGTTTATCGCGTGGGA 242
Qy 500 AGTATGTCAAGGTTTCATATSCACCGTTGCTGCTGTGGTTATGGTCTAGCTTTAT 559
Db 243 AGTACATATTTGGTTTTCGATTTGACCTAGGAGCCTCGGCCACATACACTAATTTCTCT 302
Qy 560 CCTTACAACAGCTAGCCTTTCTAAAGTCTTAAGAAGCAAAATTTCTCAGAAGTTATGG 619
Db 303 CTCTAATCGAAGTCGATTCGAGAAAGTTATTAAGAAGGAAACTTTCTCAGTCGTGTGA 362
Qy 620 ATATCATATCTACGTGAGTCTAGTGGCCAGTTGTGTTAGCGTGTGGGCTTTTGTCTA 679
Db 363 ATATCGAGATATACAGCATAGTGGCAACAGTAGCTTCTTATCGGTTTATTTGCA 422
Qy 680 CGAGTGAGTGGAAACTTTTGAGCAGTGAATGGAATACAAACATGGAAGGTATCCT 739
Db 423 CGCGGAGTGGAAAGCTTTAGAGGAGAGATGCTGCTTCAGCTCAGAGAGGTGCTCT 482
Qy 740 ACATATGAACCTAGTGTGACAGCTTTACTGSCAGTTATTTCTCATCGGTGGCACAG 799
Db 483 ATGTATGACACTTCTATGAGCTGCTGTATCTTGGCAGATAGCTTCCGTAGGAGTGTGG 542
Qy 800 GACTCATCTTCGAGCTCTCTCTATTTCTCAATGCAATAGCGTTTGGGACTCCACAG 859
Db 543 GTTTAATCTTTGTGTGTCATCACTCTTTTCAATGTGATAGCACAGTGTCTACCCA 602
Qy 860 TGGTTCCTATCTGTGCTGTATCATTTTCCATGACAAAATGAATGGGTTAAAGGTGATTT 919
Db 603 TCATTCGATTTTGTGCTGATTTTCTTCCAGCAAGATGGATGGAGTGAATATTG 662
Qy 920 CTATGATCCTAGTATTTGGGTTTCACTTCTATGCTACCAACATATCTTGATGACA 979
Db 663 CTATGTTGATGCCACTCTGGGATTCGTTTCATATGGATATCAATTATATGTCAGTGACA 722
Qy 980 AAACTTGAAGAAAAAT 996
Db 723 AGAAGGCTAGGAGACT 739

RESULT 15
LOCUS CO105310/c
DEFINITION GR_Eb0035G15.f GR_Eb Gossypium raimondii cDNA clone GR_Eb0035G15
5', mRNA sequence.
ACCESSION CO105310.1 GI:48803996
VERSION CO105310
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE 1 (bases 1 to 782)
AUTHORS Kim H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and
Wing, R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute

The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 0035 row: G column: 15.
Location/Qualifiers
1..782
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb0035G15"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/vector="pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 24.3%; Score 263; DB 7; Length 782;
Best Local Similarity 64.9%; Pred. No. 3.9e-60;
Matches 389; Conservative 0; Mismatches 210; Indels 0; Gaps 0;
Qy 387 CTCACAAAACCTTACCCCTATCAATTTAAATTTCTTTTCTTAACTATATCTTCCAC 446
Db 782 CTCACAAAAGTTACCCCTTTTCATAATAAACTCTTTGGTCTCTCTCAAACTCTCTCCAC 723
Qy 447 CCTACTTGTCAATTAATGAGGAGACAGACTCCACAAAAGTTACAAAAGGAGAGTAGTGT 506
Db 722 CCTCTCGTATTCAAAACGATTCCTCAGGTTCACACCACTCCAGAGCAAAATATGT 663
Qy 507 CAAAGGTTTCATATGACACCGTTGCTGCTGCTGCTTATGCTTAGTCTTAGTCTTATCCCTACA 566
Db 662 GATCGGATTTGTATGACGCGTGGCTGCTTCAGCAGGCTACGGATTGATGTTGTTCTCTAAC 603
Qy 567 ACAGTAGCCTTTCTAAAGTCTTAAAGAGCAAAATTTCTCAGAGTTATGGATATGAT 626
Db 602 ACACTTTGTTTCCAAAAGATTTCTCAAGAAAACAATCAATTAAGGCAGTGTGGACATGAT 543
Qy 627 AATCTAGTCTAGTCTAGTGCCAGTTGTGTAGCGTGGTGGGCTTTTGTCTAGCAGTGA 686
Db 542 AATCTACAGTATCGCTGCACTTTGATGATCAGTGTGGTCTTTTCGCGAGTGGGA 483
Qy 687 GTGAAAACCTTTGAGCAGTGAAATGGATAAATCAACAAATGGGAAGGTATCTTACATAT 746
Db 482 ATGGAAGAGTCTGGCGAAGAAATGGAAGGTTTAAACTGGGTAAGTCGGCATACGTTAA 423
Qy 747 GAACCTAGTGTGACAGCTGTACCTGGCAGTTATTTCTCCATCGGTGGCAGGACTGAT 806
Db 422 CGTTTGGTGTGATTCCTAGGTTGGCAAGTTTCTCGATTTGGTGGGTTGGAT 363
Qy 807 CTTGAGCTCTCTCTCTATTCTCAAAATGCAATAGCGTTTGGGACTCCCGAGTGGTTCC 866
Db 362 CTTGGAAGCTATCGCTTTTCTCAATGTGATCAGCACAGTGTGGACTACCCATGTTCC 303
Qy 867 TATCTTGGCTGTAATCAATTTTCCATGACAAAATGAATGGGTTAAAGGTGATTTCTATGAT 926
Db 302 AGTGTTCGCGATGTGTGTTTTTCCACGATCCAAATGACGTGAATTAAGGTCATTTCCATGTT 243
Qy 927 CTTAGCTATTTGGGTTTCACTTCTTATGCTTACCAACAATATCTTATGACAAAAACT 985
Db 242 GTTGGCCATATGGGATTTGTGCTTATGTTTACCAGCACTATCTTGTATGATGACAAACT 184

Search completed: November 1, 2004, 22:30:25
Job time : 3854.17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 06:20:04 ; Search time 618.418 Seconds
(without alignments)
9900.026 Million cell updates/sec

Title: US-09-913-767-5
Perfect score: 1194
Sequence: 1 tcatgatataataaacat.....tctgatcaagcatatttcc 1194

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 253800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCRU_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1015.8	85.1	3387	9 US-09-938-842A-2212	Sequence 2212, Ap
2	1015.8	85.1	3387	11 US-09-938-842A-2212	Sequence 2212, Ap
3	459.8	38.5	592	9 US-09-770-152-44	Sequence 44, Appl
4	349.8	29.3	1294	16 US-10-424-599-2524	Sequence 2524, Ap
5	303.8	25.4	2214	16 US-10-424-599-102519	Sequence 102519, A
6	279	23.4	1513	17 US-10-437-963-82061	Sequence 82061, A
7	247.2	20.7	997	16 US-10-424-599-75161	Sequence 75161, A
8	244.2	20.5	2381	17 US-10-437-963-43860	Sequence 43860, A
9	215.4	18.0	83698	17 US-10-416-898-9	Sequence 9, Appli
10	157.2	13.2	522	16 US-10-424-599-114931	Sequence 114931, A
11	154.4	12.9	792	16 US-10-424-599-9213	Sequence 9213, Ap
12	146.6	12.3	1386	16 US-10-425-114-12485	Sequence 12485, A

13	145	12.1	1779	16	US-10-425-114-28473	Sequence 28473, A
14	141.8	11.9	728	17	US-10-767-701-9000	Sequence 9000, Ap
15	138.8	11.6	747	17	US-10-767-701-6806	Sequence 6806, Ap
16	134.4	11.3	2605	17	US-10-437-963-83900	Sequence 83900, A
17	122.2	10.2	1146	16	US-10-425-114-10516	Sequence 10516, A
18	122.2	10.2	1233	16	US-10-424-599-30274	Sequence 30274, A
19	120	10.1	386	9	US-09-770-791-85	Sequence 85, Appl
20	118.4	9.9	332	11	US-09-732-627A-1196	Sequence 1196, Ap
21	107	9.0	1432	17	US-10-437-963-1897	Sequence 1897, Ap
22	100	8.4	671	17	US-10-767-701-5986	Sequence 5986, Ap
23	93.4	7.8	1333	16	US-10-425-114-10574	Sequence 10574, A
24	92.6	7.8	1373	16	US-10-424-599-31364	Sequence 31364, A
25	89.4	7.5	650	17	US-10-767-701-4045	Sequence 4045, Ap
26	85	7.1	327	16	US-10-424-599-121152	Sequence 121152, A
27	81.4	6.8	1355	16	US-10-424-599-31363	Sequence 31363, A
28	80.8	6.8	994	16	US-10-424-599-73501	Sequence 73501, A
29	76.8	6.4	716	16	US-10-424-599-61310	Sequence 61310, A
30	67.2	5.6	1173	17	US-10-437-963-57539	Sequence 57539, A
31	62.6	5.2	774	9	US-09-770-445-895	Sequence 895, App
32	58.2	4.9	789	17	US-10-437-963-72452	Sequence 72452, A
33	55.8	4.7	308	16	US-10-424-599-121805	Sequence 121805, A
34	55.2	4.6	1417	16	US-10-424-599-97043	Sequence 97043, A
35	52.2	4.4	1296	16	US-10-425-114-13687	Sequence 13687, A
36	49.6	4.2	737	16	US-10-425-114-24295	Sequence 24295, A
37	49	4.1	1177	17	US-10-437-963-91425	Sequence 91425, A
38	43.6	3.7	383	17	US-10-021-323-2053	Sequence 2053, Ap
39	42.8	3.6	613	17	US-10-767-701-4026	Sequence 4026, Ap
40	41.2	3.5	478	18	US-10-674-124A-2371	Sequence 2371, Ap
41	41.2	3.5	867	14	US-10-123-155-20	Sequence 20, Appl
42	41.2	3.5	867	15	US-10-146-731-20	Sequence 20, Appl
43	41.2	3.5	867	15	US-10-140-472-20	Sequence 20, Appl
44	41.2	3.5	867	15	US-10-141-761-20	Sequence 20, Appl
45	41.2	3.5	867	15	US-10-142-885-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-09-938-842A-2212

Sequence 2212, Application US/09938842A

Patent No. US20020160378A1

GENERAL INFORMATION:

APPLICANT: Harper, Jeff

APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

FILE REFERENCE: SAME, AND METHODS OF USE

FILE REFERENCE: SCRIPI300-3

CURRENT APPLICATION NUMBER: US/09/938,842A

CURRENT FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: US 60/227,866

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/300,111

PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 5379

SEQ ID NO 2212

LENGTH: 3387

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-938-842A-2212

Query Match 85.1%; Score 1015.8; DB 9; Length 3387;

Best Local Similarity 95.1%; Pred. No. 4.6e-304;

Matches 1081; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

QY 38 GTGACCAAGCTTAGAAGCAACCTTAGATCATGAGTGTGTAATCATCATCATCAT 97

DB 1205 GTGACCAAGCTTAGAAGCAACCTTAGATCATGAGTGTGTAATCATCATCATCAT 1264

```
Qy 98 CAGCTGTGCTCAAAACGAGAACTATAAAAGTGGCTTCGTGCTCCATATACGTAATCT 157
Db |||||
Qy 1265 CAGCTGTGCTCAAAACGAGAACTATAAAAGTGGCTTCGTGCTCCATATACGTAATCT 1324
Db |||||
Qy 158 TTGTCTCTTTTTCGAGCACTAGCTACAAATTTCTGGGTAGATTGTTACTATGAAATGGAG 217
Db |||||
Qy 1325 TTGTCTCTTTTTCGAGCACTAGCTACAAATTTCTGGGTAGATTGTTACTATGAAATGGAG 1384
Db |||||
Qy 218 GAAATAGACATATGTTGTAACACTTTCTTCAACTCAITGGCTTCCTGTACTGTTCTGT 277
Db |||||
Qy 1385 GAAATAGACATATGTTGTAACACTTTCTTCAACTCAITGGCTTCCTGTACTGTTCTGT 1444
Db |||||
Qy 278 TCCGCTCTTTTTCGAAATCAGCAACCAAAATCAACAGATACAAATTTCAAGTCAGTCCC 337
Db |||||
Qy 1445 TCCGCTCTTTTTCGAAATCAGCAACCAAAATCAACAGATACAAATTTCAAGTCAGTCCC 1504
Db |||||
Qy 338 CTTCTCTTCAACCACTTGCATCGGTTTACTTGTGCACCTGGACTGCTAGTCCGCTTATG 397
Db |||||
Qy 1505 CTTCTCTTCAACCACTTGCATCGGTTTACTTGTGCACCTGGACTGCTAGTCCGCTTATG 1564
Db |||||
Qy 398 CTTATTGTGTGCACTAGGTTGCTCTACTTACCAAGTCTCTACTTTCTCCCTCATCTTGG 457
Db |||||
Qy 1565 CTTATTGTCTGCA----- 1578
Db |||||
Qy 458 CCTCACAGTTGGCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTC 517
Db |||||
Qy 1579 -----TTGGCTCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTC 1630
Db |||||
Qy 518 CTTTGATAGTCAGTTCTTTTGTCTTCTTCCCTCACTGTATCTCTGCTCTTCTTGTGGTCAACA 577
Db |||||
Qy 1631 CTTTGATAGTCAGTTCTTTTGTCTTCTTCCCTCACTGTATCTCTGCTCTTCTTGTGGTCAACA 1690
Db |||||
Qy 578 CTGATTTCAGAAACTCAACTAATGATCTAGAGTACAGTATGTGATCGGGTTTCATATGTA 637
Db |||||
Qy 1691 CTGATTTCAGAAACTCAACTAATGATCTAGAGTACAGTATGTGATCGGGTTTCATATGTA 1750
Db |||||
Qy 638 CCATCGGTGCTTCCGCTGGATTGACATGTTACTATCTCTGATACAAATGCTCTTCAGGA 697
Db |||||
Qy 1751 CCATCGGTGCTTCCGCTGGATTGACATGTTACTATCTCTGATACAAATGCTCTTCAGGA 1810
Db |||||
Qy 698 AAGTTTTCACGAAGCATACCTCCTCAGCAGTCACGGACTTGGCCATTTTACCAGTCTCTAG 757
Db |||||
Qy 1811 AAGTTTTCACGAAGCATACATCTCTCAGCAGTCACGGACTTGGCCATTTTACCAGTCTCTAG 1870
Db |||||
Qy 758 TTGCAGTTGTGTAGTTCTCATAGACATTTTGGAAAGTGGAGATGGGAAACTTTGGCAA 817
Db |||||
Qy 1871 TTGCAGTTGTGTAGTTCTCATAGACATTTTGGAAAGTGGAGATGGGAAACTTTGGCAA 1930
Db |||||
Qy 818 GTGAGATGAGAACTACAACTCGGGAAGTGTATATGTTTGAATTTAGCCTTCGGCAG 877
Db |||||
Qy 1931 GTGAGATGAGAACTACAACTCGGGAAGTGTATATGTTTGAATTTAGCCTTCGGCAG 1990
Db |||||
Qy 878 CTTATTCTTGGCAAGTCTACACTCTCTGCTTGTGGGATTTGATCTTCAGTCACTCTCTG 937
Db |||||
Qy 1991 CTTATTCTTGGCAAGTCTACACTCTTGTGCTTGTGGGATTTGATCTTCAGTCACTCTCTG 2050
Db |||||
Qy 938 TGTCTCCAAATTCATACAGCTGTGGGATTTGCCCTATATAGTTCCAGTTCGGGAGTGATAG 997
Db |||||
Qy 2051 TGTCTCCAAATTCATACAGCTGTGGGATTTGCCCTATATAGTTCCAGTTCGGGAGTGATAG 2110
Db |||||
Qy 998 TTTTCCATGATAGAAATGGAAGTATGCAAAATCTTCTCCATTTTGTAGCTATCTCGGCT 1057
Db |||||
Qy 2111 TTTTCCATGATAGAAATGGAAGTATGCAAAATCTTCTCCATTTTGTAGCTATCTCGGCT 2170
Db |||||
Qy 1058 TCTTTTCAATTCGTCTATCAGCACTACCTCGAGAAAGAAAGTTGAATTAAGTCAAGCAACA 1117
Db |||||
Qy 2171 TCTTTTCAATTCGTCTATCAGCACTACCTCGAGAAAGAAAGTTGAATTAAGTCAAGCAACA 2230
Db |||||
Qy 1118 GTGCTGTAGAGATCTTCACTTACCTGTGTAGGAAAGTTCACAAACATACAAAGTG 1174
Db |||||
Qy 2231 GTGCTGTAGGAGATCTTCACTTACCTGTGTAGGAAAGTTCACAAACATACAAAGTG 2287
Db |||||
```

```
RESULT 2
US-09-938-842A-2212
; Sequence 2212, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krops, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2212
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2212

Query Match      85.1%; Score 1015.8; DB 11; Length 3387;
Best Local Similarity 95.1%; Pred. No. 4.6e-304;
Matches 1081; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

Qy 38 GTGACCGAAGTCTAGAACGCAACCTTATAGATCATGAGTGGTAACATCATCATCATCAT 97
Db |||||
Qy 1205 GTGACCGAAGTCTAGAACGCAACCTTATAGATCATGAGTGGTAACATCATCATCATCAT 1264
Db |||||
Qy 98 CAGCTGTGCTTCAAAACGAGAACTATAAAAGTGGCTTCGTGCTCCATATACGTAATCT 157
Db |||||
Qy 1265 CAGCTGTGCTTCAAAACGAGAACTATAAAAGTGGCTTCGTGCTCCATATACGTAATCT 1324
Db |||||
Qy 158 TTGTCCTCTTTTGGCAGCCACTAGCTACAAATTTCTGGGTAGATTGTTACTATGAAATGGAG 217
Db |||||
Qy 1325 TTGTCCTCTTTTGGCAGCCACTAGCTACAAATTTCTGGGTAGATTGTTACTATGAAATGGAG 1384
Db |||||
Qy 218 GAAATAGACATATGTTGTAACACTTTCTTCAACTCAITGGCTTCCTGTACTGTTCTGT 277
Db |||||
Qy 1385 GAAATAGACATATGTTGTAACACTTTCTTCAACTCAITGGCTTCCTGTACTGTTCTGT 1444
Db |||||
Qy 278 TCCGCTCTTTTTCGAAATCAGCAACCAAAATCAACAGATACAAATTTCAAGTCAGTCCC 337
Db |||||
Qy 1445 TCCGCTCTTTTTCGAAATCAGCAACCAAAATCAACAGATACAAATTTCAAGTCAGTCCC 1504
Db |||||
Qy 338 CTTCTCTTCAACCACTTGCATCGGTTTACTTGTGCACCTGGACTGCTAGTCCGCTTATG 397
Db |||||
Qy 1505 CTTCTCTTCAACCACTTGCATCGGTTTACTTGTGCACCTGGACTGCTAGTCCGCTTATG 1564
Db |||||
Qy 398 CTTATTGTGTGCACTAGGTTGCTCTACTTACCAAGTCTCTACTTTCTCCCTCATCTTGG 457
Db |||||
Qy 1565 CTTATTGTGTGCA----- 1578
Db |||||
Qy 458 CCTCACAGTTGGCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTC 517
Db |||||
Qy 1579 -----TTGGCTCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTC 1630
Db |||||
Qy 518 CTTTGATAGTCAGTTCTTTTGTCTTCTTCCCTCACTGTATCTCTGCTCTTCTTGTGGTCAACA 577
Db |||||
Qy 1631 CTTTGATAGTCAGTTCTTTTGTCTTCTTCCCTCACTGTATCTCTGCTCTTCTTGTGGTCAACA 1690
Db |||||
Qy 578 CTGATTTCAGAAACTCAACTAATGATCTAGAGTACAGTATGTGATCGGGTTTCATATGTA 637
Db |||||
Qy 1691 CTGATTTCAGAAACTCAACTAATGATCTAGAGTACAGTATGTGATCGGGTTTCATATGTA 1750
Db |||||
Qy 638 CCATCGGTGCTTCCGCTGGATTGACATGTTACTATCTCTGATACAAATGCTCTTCAGGA 697
Db |||||
Qy 1751 CCATCGGTGCTTCCGCTGGATTGACATGTTACTATCTCTGATACAAATGCTCTTCAGGA 1810
Db |||||
Qy 698 AAGTTTTCACGAAGCATACCTCCTCAGCAGTCACGGACTTGGCCATTTTACCAGTCTCTAG 757
Db |||||
Qy 1811 AAGTTTTCACGAAGCATACATCTCTCAGCAGTCACGGACTTGGCCATTTTACCAGTCTCTAG 1870
Db |||||
Qy 758 TTGCAGTTGTGTAGTTCTCATAGACATTTTGGAAAGTGGAGATGGGAAACTTTGGCAA 817
Db |||||
Qy 1871 TTGCAGTTGTGTAGTTCTCATAGACATTTTGGAAAGTGGAGATGGGAAACTTTGGCAA 1930
Db |||||
Qy 818 GTGAGATGAGAACTACAACTCGGGAAGTGTATATGTTTGAATTTAGCCTTCGGCAG 877
Db |||||
Qy 1931 GTGAGATGAGAACTACAACTCGGGAAGTGTATATGTTTGAATTTAGCCTTCGGCAG 1990
Db |||||
Qy 878 CTTATTCTTGGCAAGTCTACACTCTCTGCTTGTGGGATTTGATCTTCAGTCACTCTCTG 937
Db |||||
Qy 1991 CTTATTCTTGGCAAGTCTACACTCTTGTGCTTGTGGGATTTGATCTTCAGTCACTCTCTG 2050
Db |||||
Qy 938 TGTCTCCAAATTCATACAGCTGTGGGATTTGCCCTATATAGTTCCAGTTCGGGAGTGATAG 997
Db |||||
Qy 2051 TGTCTCCAAATTCATACAGCTGTGGGATTTGCCCTATATAGTTCCAGTTCGGGAGTGATAG 2110
Db |||||
Qy 998 TTTTCCATGATAGAAATGGAAGTATGCAAAATCTTCTCCATTTTGTAGCTATCTCGGCT 1057
Db |||||
Qy 2111 TTTTCCATGATAGAAATGGAAGTATGCAAAATCTTCTCCATTTTGTAGCTATCTCGGCT 2170
Db |||||
Qy 1058 TCTTTTCAATTCGTCTATCAGCACTACCTCGAGAAAGAAAGTTGAATTAAGTCAAGCAACA 1117
Db |||||
Qy 2171 TCTTTTCAATTCGTCTATCAGCACTACCTCGAGAAAGAAAGTTGAATTAAGTCAAGCAACA 2230
Db |||||
Qy 1118 GTGCTGTAGAGATCTTCACTTACCTGTGTAGGAAAGTTCACAAACATACAAAGTG 1174
Db |||||
Qy 2231 GTGCTGTAGGAGATCTTCACTTACCTGTGTAGGAAAGTTCACAAACATACAAAGTG 2287
Db |||||
```

Db 1751 CCATCGGTGCTCCGCTGGGATTTGGACTGTATTCTATCTCTGATACAAAATGCTCTTCAGGA 1810
QY 698 AAGTTTTTCAGGAAGCATACATCCTCAGCAGTCAAGGACTTTGGCCATTTTACCACTCTCTAG 757
Db 1811 AAGTTTTTCAGGAAGCATACATCCTCAGCAGTCAAGGACTTTGGCCATTTTACCACTCTCTAG 1870
QY 758 TTCCGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAAGTGGAGAGTGGGAAACTTTTGCCAA 817
Db 1871 TTCCGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAAGTGGAGAGTGGGAAACTTTTGCCAA 1930
QY 818 GTGAGATGAGAAACTCAAACTCGGAAAGTGTTCATATGTTTGTAGCTTTAGCTCGGCAG 877
Db 1931 GTGAGATGAGAAACTCAAACTCGGAAAGTGTTCATATGTTTGTAGCTTTAGCTCGGCAG 1990
QY 878 CTATTTTCTGGCAAGTCTACACTCCTCGTCTTGTGGGATTTGATCTTCGAGTCACTCTCTG 937
Db 1991 CTATTTTCTGGCAAGTCTACACTCCTCGTCTTGTGGGATTTGATCTTCGAGTCACTCTCTG 2050
QY 938 TGTCTTCCAAATTCATAACAGCTGTGGGATTTGCTATAGTTCCAGTTGGCGAGTGATAG 997
Db 2051 TGTCTTCCAAATTCATAACAGCTGTGGGATTTGCTATAGTTCCAGTTGGCGAGTGATAG 2110
QY 998 TTTTTCATGATGAATGGAGGATCCAAATCTTCTCCATTATTTTAGCTATCTCGGCT 1057
Db 2111 TTTTTCATGATGAATGGAGGATCCAAATCTTCTCCATTATTTTAGCTATCTCGGCT 2170
QY 1058 TCCTTTTCAATTCGCTATCAGCACTACCTCGACGAAAGAGTTGAATCTAGCCACACAA 1117
Db 2171 TCCTTTTCAATTCGCTATCAGCACTACCTCGACGAAAGAGTTGAATCTAGCCACACAA 2230
QY 1118 GTGCTGTAGAGATCTTCATCTACCTGTGGAGAGGTCAACAAACATACAAAGTG 1174
Db 2231 GTGCTGTAGAGATCTTCATCTACCTGTGGAGAGGTCAACAAACATACAAAGTG 2287

RESULT 3

US-09-770-152-44
; Sequence 44, Application US/09770152
; Publication No. US20020040489A1
; GENERAL INFORMATION:
; APPLICANT: Grolach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2025US (PARA-014PRV)
; CURRENT APPLICATION NUMBER: US/09/770,152
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,503
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-152-44

Query Match 38.5%; Score 459.8; DB 9; Length 592;
Best Local Similarity 91.6%; Pred. No. 9.4e-132;
Matches 501; Conservative 0; Mismatches 37; Indels 9; Gaps 1;
QY 37 GGTGACCGAAGCTTAGAAGCAAACTTATAGATCATAGAGGTGGTAACCTGAATCATCATCA 96
Db 55 GGTGACCGAAGCTTAGAAGCAAACTTATAGATCATAGAGGTGGTAACCTGAATCATCATCA 105
QY 97 TCAGCTGTGCTCAAAACCGAGAACTATAAAGGTGGTTCGTTGCTCCATATACGTAATC 156
Db 106 TTTTCAGTACCTCAAAACCGAGAACTGTAGAGGTGGTTCGTTGCTCCATATACGTAATC 165
QY 157 TTTTGTCTCTTTTGGCAGCCACTAGCTACAATTTCTGGGTAGATTTGATCTATGAAATGGA 216
Db 166 TTTTGTCACTCTTCTGCCAACCACTTGTCTACAGTTCTGGGTAGACTGTACTATGAAATGGA 225
QY 217 GGAATAGACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTGTTCTG 276
Db 226 GGGAAAAGCACATATGTGGTAACACTTCTTCAACTCATTTGGCTTCCCTGTACTGTATCTC 285
QY 277 TTTCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC 336
Db 286 TTTCCGCTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC 345
QY 337 CCTTCTTCAACACCCCTTGCATCGGTTTACTTTGTGCACCTGGACTGTAGTCCGCTTAT 396
Db 346 CCTTCTTCAACACCCCTTGCATCGGTTTACTTTGTGCACCTGGACTGTAGTGTCTGCTTAT 405
QY 397 GCTTATTTGTCTGAGTAGGGTTGCTCTACTTACAGTCTCTACTTTTCTCCCTCATCTTTG 456
Db 406 GCTTATTTGTCTGAGTTGGTTGCTTACTTACCAAGTCTCTACTTTTCTCCCTCATCTTTG 465
QY 457 GCTTCACAGTTGGCCCTTCACTGCCCTTTTCTCATATTTTCCCTAACTCGCAAAAGTTCACT 516
Db 466 GCTTCACAGTTGGCCCTTCACTGCCCTTTTCTCATATTTTCTTAATTCGCAAAAGTTCACT 525
QY 517 CCTTTGTAGTACAGTTCTTTTGTCTTCTCTCACTGTATCTCTGTCTTCTTGTGGTCAAC 576
Db 526 CCTTTGTAGTACAGTTCTTTTGTCTTCTCTAGGGTTTCTCTGCTTCTCTGCTGCTCAAC 585
QY 577 ACTGATT 583
Db 586 ACTGATT 592

RESULT 4
US-10-424-599-2524
; Sequence 2524, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 2524
; LENGTH: 1294
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_102285C.1
US-10-424-599-2524

Query Match 29.3%; Score 349.8; DB 16; Length 1294;
Best Local Similarity 59.5%; Pred. No. 2.2e-97;
Matches 622; Conservative 0; Mismatches 402; Indels 21; Gaps 1;
US-10-424-599-2524

Db 1020 ATCAATTCGTGCTCTGGTCTTCTGTCTATAGCCCTATTTTGCAGTGGGAATGGCGTAC 1079
Qy 810 TTTCGAAGTGAGATGAGAACTACAAACTCGGAAAGTGTCTATATGTTTTCGACTTTAGC 869
Db 1080 TTTCGATGGAGAATGGAGGTTTTCAGAAAGGATATGTTGCTTAATGTTATGCTTTCGT 1139
Qy 870 CTCGGCAGCTATTTCTGCGAAGTCTACACTCTGCTCTGTTGGGATTTGATCTTCGAGTC 929
Db 1140 TTGGACTTCAATAGCCTGGCAGGTATGCTCTGTTGGTGTGTTGGCTTTCCTAGT 1199
Qy 930 ATCTCTGTGTTCTCCAAATTCATAACAGCTGTGGGATTTGCCCTATAGTTCCAGTTTCGGC 989
Db 1200 GTCTCTCTACTCTCAATGTTTAAAGCAGTTTCTTAGCCGTAACCTCTATGCTGC 1259
Qy 990 AGTGATAGTTTTCATGATAGATGACGCATCCAAAATCTTCTCCATTATTTTAGCTAT 1049
Db 1260 TGTATAGTTTTTCATGATAAGATGAATGGGTGAAGATAATTTCTATGCTTTTGGCTCT 1319
Qy 1050 CTGCGGCTTCCTTTCATTCTGCTATCAGCACTACCTCGAGCAAAAGAAG 1098
Db 1320 ATGGGGTTTTGCTCTTATATTTATCAGAAATTTATCTTGATGATTTCAAAG 1368

RESULT 6

US-10-437-963-82061/c

; Sequence 82061, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yihwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 82061

; LENGTH: 1513

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_81525C.1

US-10-437-963-82061

Query Match 23.4%; Score 279; DB 17; Length 1513;
Best Local Similarity 55.8%; Pred. No. 2.4e-75;
Matches 556; Conservative 0; Mismatches 435; Indels 6; Gaps 1;
Qy 107 CTCGAACCGAGAACTATAAAGTGGCTTCGTCTCCATATACGTAATCTTTTGGCTCT 166
Db 1278 CTCTCTCTCAAGAGTTGGCAATGGTGGCTGATGGTGGAGTTAACATGTTCTTCTCATTTG 1219
Qy 167 TTTCGAGCCATAGCTACAAATCTGGGTAGATTGCTATGAAATGGAGGAATAGCA 226
Db 1218 CCGGTCAACCGCATCTACACTTCTTGGGAGGTTCTACTACAAATCAAGCGCGCAATAGCA 1159
Qy 227 CATATGTGTAACACTTCTTCAACTCATTTGGCTTCCCTGCTACTGTTCTGTTCCGCTTCT 286
Db 1158 AGTGGATGTCACATTCGTCGAACCTGGCTTTCGAGTTTGTTCATGTCCTTATTC 1099
Qy 287 TTTCGGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCCCC-----TT 340
Db 1098 TTTTTCATTCAAAGACATCTTCTACACAAACAGTCACTAGTAGTCTGCTGCCCTACAAATTT 1039
Qy 341 CTTTCAACACCTTGCATGGTTTACTTGTGCAGTGGACTGTAGTGTCCGCTTATGCTT 400
Db 1038 CTATCCCCAAAATTACTCTGATATATGTTGTCTCGGCGCTCATCATTTGCTGCAGACGACT 979

Qy 401 ATTTGTCTCAGTAGGTTGCTCTACTTACCAGTCTCTACTTTTCTCCCTCACTCTTGGCCT 460
Db 978 TGATGTAATTCCTATGCGCTACTATATCTCCGGTCTCAACATATTCGCTCACTGTGCTA 919
Qy 461 CACAGTTGGCCTTCACTGGCTTTTCTCATATTTCTTAACTCGGAAAGTTCACTTCCTT 520
Db 918 GTCACTTGGCCTTCAATGCTGTCTTCTCATATTTCTCAATGCTCAAAAATTCACCCCTC 859
Qy 521 TGATAGTCAGTTCTTTCCTTCTCCCTCACTGTATCTCTGCTCTTCTTGTGGTCAACACTG 580
Db 858 TGATTTTCAATTCCTGATGCTCTTCTAGTTTCTGCTTCACTCTCTGGAGTTGATGAAG 799
Qy 581 ATTCAGAAAACTCAACTAATGTATCTAGAGTACAGTATGTGATCGGGTTTCATATGTACCA 640
Db 798 ATTCCTCAGGAACACTACTAGTATATACATGGAAGTACATTTTGGGTTTCTCTGTTGACAC 739
Qy 641 TCGGTGCTTCGCTGGGATTTGGACTGTACTATCTCTCATACAAATGTCTTTCAGAAAG 700
Db 738 TAGGGGCTATCAGCTACATCTCGCTCATTTCTCCCTGATGCAAGTCAATTTGAGAAG 679
Qy 701 TTTTCACGAAGCATACATCTCAGCAGTCACGGACTTGGCCATTTACCAAGTCTCTAGTTG 760
Db 678 TTATTAAGAGGAGACCTTCTCAGTTGTGTGAACATGCAGATATATACAGCTCTCTGCTG 619
Qy 761 CGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAGTGGAGAGTGGGAAACTTTTGCCAAAGTG 820
Db 618 CAACATTGGCTTCTCTTCTGTTGGTTATTTTTCGAAGTGGTGAATGGATGACTTTTACAAGGAG 559
Qy 821 AGATCAGAAAACTACAAACTCGGAAAGTGTCTATGTTTGTGACTTTAGCTTCGGCAGCTA 880
Db 558 AGATGCATGCATTTCCAGTCTGGGAAAGTGTCTATATGTAATGACACTGTGTGGACGGCTA 499
Qy 881 TTTTCCTGCAAGTCTPACACTCTGCTGTCTTGTGGGATTTGATCTTCCAGTCACTCTCTGCT 940
Db 498 TATCTTGGCAGTAGCATCAGTTGGAGTGGTGGGATTTGATCTTCTGTTGTTGCTCATGCTGT 439
Qy 941 TCTCAATTTCCATAACAGCTGTGGGATTTGCCCTATAGTTCCAGTTCGGCGCAGTGATAGTTT 1000
Db 438 TTTCAATGTGTATAAGCACCCCTAGCTTACCCATCATCTCTGTTTTCGTGTGATTTTCT 379
Qy 1001 TCCATGATAGAAATGGACGCATCCAAAATCTTCTCCATTTATTTTAGCTATCTCGCGCTCC 1060
Db 378 TTCATGACAAGATGGATGGAGTAAAGATTATAGCTATGCTGATGGCCATTTGGGATTTA 319
Qy 1061 TTTTCATTCGTCTATCAGCACTACCTCGACGAAAGAA 1097
Db 318 TGTCTATATGGCCACCACCAATATATGTTGATGGCAAGAA 282

RESULT 7

US-10-424-599-75161

; Sequence 75161, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 75161

; LENGTH: 997

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_38884C.1

US-10-424-599-75161


```
Query Match      20.7%; Score 247.2; DB 16; Length 997;
Best Local Similarity 57.3%; Pred. No. 1.4e-65;
Matches 500; Conservative 0; Mismatches 348; Indels 24; Gaps 2;

Qy 118 AACTATAAAGGTGGCTTCGTGTCTCCATATACGTAATCTTTGTGCTCTTTTGGCAGCCA 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 125 AAATACAAATGGTGCTACGGTGTGCTCTACATAATCTTTCTTTGTAGGCCAATCT 184

Qy 178 CTAGCTACAATCTGGGTAGATTTGATATGAAATGGAGAAATAGACATATGTGGTA 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 CGAGCACTCTTTTGGGAAGATTACTATGACAAATGGTGTGTAATAGCAATGGATGCA 244

Qy 238 ACACCTCTTCAACTATGGCTTCCCTGTAATGTTGTTCCGCTTCTTTCTCGA--- 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 ACAITTTGTTCAATCAGCTAGGATTCCTGTACTTCTTCCACCTCTTTTACTTCCCA 304

Qy 295 -----ATCAGGCAACCCCAATCAACAGATACAAAATTTTCAGTCAGTCC 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 305 CAACACATGCCAAGTTTAAATAACAATCCAGCAATATGATTTCTACAAAACAAA 364

Qy 337 CTTCTCTTACACACCCTTGGCATCGGTTTACTTTGTGCACTGGACTGCTAGTGTCCGTTAT 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 365 CCAAAATTTCTCCACCCTTGTCTCTCTATCTAGCTTTTGGCCTAAATTTCTCACAGGGAC 424

Qy 397 GCTTATTTGTCTGCACTAGGTTGCTCTACTTACCAGTCTCTACTTTCTCCCTCATCTTG 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 425 AACTTGAATTTCTTATGGACTTTTATATCTCCCTCTTTCCACCTATTTCTTACTATGT 484

Qy 457 GCCTCACAGTTGGCCCTTCACTGCCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 485 GCCACCAATAGTCTTCAACGGGTGTTCTTTCTTCTCAATTTCCAGAAATTCACA 544

Qy 517 CTTTGTAGTCAAGTTCTTGTCTCTCTACTGATCTCTCTGCTTCTTCTTGTGGTCAAC 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 545 GCATTTCAATCAATTTCTGTAGTCTCTTACCATATCAGCTTCTCTGCTTGCATCAAC 604

Qy 577 ACTGATTC--AGAAAACTCAACTAAATGATCTAGATACAGTATGTATCGGGTTTATA 633
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 605 TCCGATTCGAGAGGACTCAACCGGTCTTTCCAGAGAAAGCATGTAATTTGGATTTCT 664

Qy 634 TGTACATCGGTGCTTCGCTGGGATTTGGAATGTTACTATCTCTGATACAAATGCTTTC 693
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 665 TGCACCATTTGGTCACTCAGTACATCTCTTTGTACCTCTCTCTGTGACGCTTCTTTT 724

Qy 694 AGGAAAGTTTCAAGAGCATATATCTTCAGAGTCAAGGACTTGGCCCATTTACAGTCT 753
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 CAGAAAGTTATAAGAGAGAAACCTTTTCTGCTGTGTTGGACATGCAATTTTACCCATCC 784

Qy 754 CTAGTTGCGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAGTGGAGAGTGGAAACTTGG 813
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 TTCAATGCTCATGTGCTTGTGTAGTAGGGTTGTTTTCGAAGTGGAGAGTGGAAAGTTTG 844

Qy 814 CCAAGTGAATGAGAAACTACAACTCGGGAAGTGTATATGTTTGTGACTTTTACGCTCG 873
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 AACATGAGATGAAGGATATGATAAGGAAGTGTGCCCGTTGTAATGACTGTACTATAG 904

Qy 874 GCAGCTATTTCTGCAAGTCTACACTCTCTGCTTGTGGGATGTATCTTCGAGTCATCC 933
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 ATTGCTGTGACATGTATATATGTTCTACTAGGGAATGCTAGGGGGACGTTTGGAGAGTCT 964

Qy 934 TCTGTGTTCTCAATTTCCATAACAGCTGTGGG 965
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 965 GCATTTGTCGGATGTGAGAAGTACCTGTTGGG 996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 8
US-10-437-963-43860/c
; Sequence 43860, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

```
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 43860  
; LENGTH: 2381  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(2381)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT_MRT4530_46977C.1  
US-10-437-963-43860
```

Query Match 20.5%; Score 244.2; DB 17; Length 2381;
Best Local Similarity 53.5%; Pred. No. 2e-64;
Matches 544; Conservative 0; Mismatches 458; Indels 15; Gaps 1;

```
Qy 89 CATCATCATCAGCTGTCCTCAAAACCGAAGAACTATAAAAGTGGCTTCGTGTCTCCATAT 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2284 CAGCAGCTCCGGAAACACACAGGCAAGCAATTCGATGTTGGCGATGTTGGCAGTGG 2225

Qy 149 ACGTAATCTTTGTCTCTTTTGGCAGCCACTAGTACAAATCTTGGGTAGATTGTACTATG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2224 ACGTCTTCTTCTCTCATCGCGGTGAGACATCGGCGACACTGCTGGGAGATACTACTACA 2165

Qy 209 AAAATGGAGGAAATAGCACATATGTGTTAAACACTTCTTCAACTCATTTGGCTTCCCTGTAC 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2164 CCCAAGCGCGCCGACGAGTGGATATCGGGTTTCGTGCGGACGGCCGGCTTCCCGATAC 2105

Qy 269 TGGTCTTGTTCGGCTTCTTTTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCA 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2104 TGTCTTTCACCTCTTCTTCTTCCCTCGAAGTCACCGTCTC-----CT 2060

Qy 329 GTGAGTCCCTTCTTCCACACCTTGCATCGGTTTACTTGTGCACTGGAAGTGTAGTGT 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2059 GCACCAACACTCCCATGGCCAAAGCTCGCGGTGATATACATCGTCTGGGCTCATCATCG 2000

Qy 389 CCGCTTATGCTTATTTCTGTCAGTAGGGTTGCTCTACTTACCAGTCTCTACTTTCTCCC 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1999 CCGCGAGCAGCATGATGTACACTGGTGGCTCAAGTACCTCCCGCTTCGACCTACTCGC 1940

Qy 449 TCATCTTGGGCTCACAGTTGGCTTCACTGCTTTTTTCTCATATTTCTTAACTTCGCAAA 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1939 TCATCTGTCAGCAGCTCGCGTTCAATGTCGTCTTCTCATAGTGTCTCAACTCCGAGA 1880

Qy 509 AGTTCACTCTTTGATAGTCACTTGTGTTCTCTCCTCACTGTATCTCTCTCTCTCTTG 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1879 AGGTCATCTCTCTGATATTCAACTCCGTGCTGCTGCTCAACATGTCGCTTCGCTCATCG 1820

Qy 569 TGGTCAACACTGATTCAGAAAACTCAACTAATGTATCTAGAGTACAGTATGTATCGGGT 628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1819 GAGTCAGCAAGGAGTCTCAGGGGGTCAACCGGCTCTCGGGAGGGAAGTATCTGCTCGTT 1760

Qy 629 TCATATGTACATCGGTGCTTCCGCTGGGATTTGAGCTGTTACTATCTCTGTATACAAATGC 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1759 TCGTGTGAGCTGGGGCGTCTGTGCACCTACTCGCTGATCTCTCGCTCATGAGGTCA 1700

Qy 689 TCTTCAGGAAAGTTTTCACGAAGCATACATCTCAGCAGTCAACGACTTGGCCATTAC 748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1699 CTTTCGAGAGCCATCATCAAGAGCACACCTTCTCAGCCGCTCTCAACATGCAGATCTACA 1640

Qy 749 AGTCTCTAGTTGCGAGTTGTGTAGTCTCTCATAGACTTTTGTGCAAGTGGAGAGTGGGAAA 808
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Db 1639 CGCGGCTCGTGGCGACGGCCGCGTCCGTGTCGGGCTGTTCCGAGCGGCGAGTGAGGT 1580
QY 809 CTTTGGCAAGTGAGATGAGAAATACAAACTCGGGAAGTGTCTATATGTTTTCGACTTTAG 868
Db 1579 CGCTGAGGGGGAGATGAACGCTTCAGGTTCGGGCGAGTTCCTACCTGATGACGCTGC 1520
QY 869 CCTCGGCGAGTATTTCTGCGCAAGTCTACACTCCTGGTCTGTGGGATGATCTTCGAGT 928
Db 1519 TGTGGGCGGCGGTGTCGTGGCAGGTGGCCAACTCGGGGTGCTCGGCCCTCATCTTCGAGG 1460
QY 929 CATCTCTGTGTTCTCCAAATTCATAACAGCTGTGGATGGCTATAGTTCAGTTGCGG 988
Db 1459 TGTGCGGGCTCTCTCCAGCTGATCAGACGGTGTCTCGCGGCTATCCGCTTCTTCG 1400
QY 989 CAGTGATAGTTCATGATAGATGAGATGAGCATCCAAATCTCTCCATTTATTTAGCTA 1048
Db 1399 CGGTGCTGTGTTCCACGACGAGTGAACGGGGTGAAGATCGTGGCCATGCTGATGCAA 1340
QY 1049 TCTGCGGCTTCCTTTTCATTCGTCTATCAGACATCTACCTCGACGAAAGAGTTGAATA 1105
Db 1339 TTTGGGGATTTATTTCTGATCTCTTTTCAGCACTATCTAGATGGCAAGAAAGCAAGA 1283

RESULT 9

US-10-416-898-9
; Sequence 9, Application US/10416898
; Publication No. US20040172670A1
; GENERAL INFORMATION:
; APPLICANT: Yale University
; APPLICANT: Walker, Elsbeth
; APPLICANT: Dellaportia, Stephen
; TITLE OF INVENTION: MAIZE YELLOW STRIPE1 AND RELATED GENES
; FILE REFERENCE: 44574-5106-US
; CURRENT APPLICATION NUMBER: US/10/416,898
; CURRENT FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: PCT/US01/43101
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 60/249,222
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 83698
; TYPE: DNA
; ORGANISM: Zea mays
US-10-416-898-9

Query Match 18.0%; Score 215.4; DB 17; Length 83698;
Best Local Similarity 52.2%; Pred. No. 1.3e-54;
Matches 508; Conservative 0; Mismatches 456; Indels 9; Gaps 1;
QY 125 AAAGGTGGCTTCGTGTCCTCATATAGTAATCTTTCTGCTCTTTGCGAGCCACTAGCTA 184
Db 67226 AATGGTGGATCTGTGTTTCAATAGCAATCTTTCTCATCTCTGCTCAAGCAATTCG 67285
QY 185 CAATTCGGGTAGTGTACTATGAAATCGAGAAATAGCACATATGTGTAACACTTC 244
Db 67286 TTCTTCTGTCGGTTTTATTACAAGAGGTGGAACAGTAATGATCTCTACTCTTG 67345
QY 245 TTCAACTCATTTGCTCCCTGACTGCTGTTCTGCTGCTCTTTTCTCGAATCAGGCAAC 304
Db 67346 TTCAAACTGGTGGCTTCCCAATCTTTATCTCCTCTTTCTCTCTCTCTCTCTCAAT 67405
QY 305 CCAAAATCAACAGATACAAATTTTCAGTCAGTCCCTTCTTCCACCCTTGCATCGTTT 364
Db 67406 CTTCTTCTTCT 67456
QY 365 ACTTGTGCACTGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 424
Db 67457 ATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 67516
QY 425 ACTTACCACTGCTACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 484

Db 67517 ATCTCTCTCTCTCAACTTATTCGATTTCTATGTGCTTTCACAGTTAGCTTTCAATGCTGTCT 67576
QY 485 TCTCATATTTCTTAACTCGCAAAAGTTTCACTCTCTTTGATAGTACAGTCTCTTTCTCTCTCC 544
Db 67577 TCTATTTATACATCAATCTCAGAAATCACTGTTGATTTTCTTCTCAGTGTGTTTTC 67636
QY 545 TCACGTATCTCTCTGCTCTTCTGTTGGTCAACACTGATTCAGAAACTCAACTCAATTAATGTAT 604
Db 67637 TCTCTATCTCTGCTGTTGTTGTTTCTCTTGAAGATTTCAATAGCCCATCAGGAGATT 67696
QY 605 CTAGAGTACAGTATGTGATCGGGTTCATATGTACCATCGGTGCTTCGGTGGGATGGAC 664
Db 67697 CTAAGTGGAGTTACTTGTGATTTGGTGTCTGTGCAAGTTTGTCTCTCTTATCTATCTCTC 67756
QY 665 TGTACTATCTCTGATACAAATGCTCTTCAGGAAAGTTTTCAGAAAGATACATCTCTCAG 724
Db 67757 TTCAGTCTCTCTTATGCAAGTTTCTTTCAGAAAGTTTCTCAAGTGTAGACTCTCTCTA 67816
QY 725 CAGTCAAGGACTTGGCCATTTTACAGTCTCTTGTGTCGAGTGTGTGTAGTTCATAGGAC 784
Db 67817 TGTGTTCTCGAGATGCAAAATCTATACGTGCTTGTGCTTCTTGTGTAGCGTTATCGGAT 67876
QY 785 TTTTTCAGAGTGGAGTGGGAAACTTTCGCAAGTGTGATGAGATGAGAAACTCAAACTCGGA 844
Db 67877 TGTTCGCAAGCGGGGAATGGATGTTGTTGAGTGTGAGATGGAAGAGTTTTCAGGAAGGTC 67936
QY 845 AAGTGTCTATGCTTTTGTAGCTTTTAGCTCGGAGTATTTCTTGGCAAGTCTACACTCTCTG 904
Db 67937 AAGTCATTTATGTTTGTAGCTTTGTTGGGCGAGCGGTTTCTGTTCAATTTGGTGTGTAG 67996
QY 905 GTCTTGTGGGATTTGATCTTCGAGTCTATCTCTGTGTTCTTCCAAATTCATAACAGCTGTGG 964
Db 67997 GAGCGGTGTCGCTTATTTCTGCTGCTCTTCTGCTGTTTCAAACCTTATTAGTACGCTCT 68056
QY 965 GATTCGCTATAGTTCAGTTCGGGCGAGTATGATTTTCCATGATAGATGGAGCATCCA 1024
Db 68057 CACTCATTTGTTACGCTCTCTCGCGGCAATTCGCGGTGTTCCATGACAGCTGACTGAGTTA 68116
QY 1025 AAATCTTCTCCATTTATTTTAGTCTATCTGCGGCTTCTCTTTCATTCGTCATCAGCACTACC 1084
Db 68117 AGATGTTGGGATGCGCATCGCTTCACAGGATTCAGTTTATATCTACCAAGCACTATC 68176
QY 1085 TCGACGAAAGAA 1097
Db 68177 TTGATGACTTGAA 68189

RESULT 10

US-10-424-599-114931
; Sequence 114931, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 114931
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_74793C.1
US-10-424-599-114931

Query Match 13.2%; Score 157.2; DB 16; Length 522;
Best Local Similarity 58.0%; Pred. No. 9e-38;
Matches 296; Conservative 0; Mismatches 213; Indels 1; Gaps 1;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:48:03 ; Search time 110.151 Seconds
(without alignments)
7704.726 Million cell updates/sec

Title: US-09-913-767-5
Perfect score: 1194
Sequence: 1 tcatgagataataaacat.....tgtgatcaagcatatttcc 1194

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCBUS COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80.4	6.7	7218	1	US-08-232-463-14
2	45	3.8	1141	4	US-09-806-708B-22
3	41.6	3.5	289	3	US-09-007-005-17
4	41.6	3.5	289	3	US-09-244-796-17
5	37	3.1	399	4	US-09-621-976-8976
6	36.8	3.1	832	4	US-09-621-976-2813
7	35.8	3.0	1141	4	US-09-806-708B-22
8	35.2	2.9	248	3	US-09-007-005-32
9	35.2	2.9	248	3	US-09-244-796-32
10	35.2	2.9	277	3	US-09-007-005-3
11	35.2	2.9	277	3	US-09-244-796-3
12	34.8	2.9	1923	4	US-09-134-000C-466
13	34.4	2.9	832	4	US-09-621-976-2813
14	34.4	2.9	1318	4	US-09-719-108-1
15	34.4	2.9	1830121	4	US-09-557-884-1
16	34.4	2.9	1830121	4	US-09-643-990A-1
17	34.4	2.9	1830121	4	US-10-329-960-1
18	34.2	2.9	474	4	US-09-621-976-18033
19	34.2	2.9	591	4	US-09-495-050A-104
20	34.2	2.9	2412	4	US-09-620-312D-1023
21	34.2	2.9	4211	3	US-09-004-838-106
22	33.8	2.8	135	4	US-09-513-999C-33549
23	33.8	2.8	1860	4	US-09-919-497-50
24	33.6	2.8	282	3	US-09-461-697-205
25	33.6	2.8	306	3	US-09-461-697-203
26	33.6	2.8	696	3	US-09-461-697-193
27	33.6	2.8	699	3	US-09-461-697-191

C 28	33.6	2.8	717	3	US-09-461-697-189	Sequence 189, App
C 29	33.6	2.8	774	3	US-09-461-697-187	Sequence 187, App
C 30	33.6	2.8	801	3	US-09-459-956-7	Sequence 7, Appli
C 31	33.6	2.8	819	3	US-09-461-697-185	Sequence 185, App
C 32	33.6	2.8	1669	3	US-09-461-697-184	Sequence 184, App
C 33	33.6	2.8	4026	4	US-09-248-796A-4971	Sequence 4971, Ap
C 34	33.2	2.8	1830121	4	US-09-557-884-1	Sequence 1, Appli
C 35	33.2	2.8	1830121	4	US-09-643-990A-1	Sequence 1, Appli
C 36	33.2	2.8	1830121	4	US-10-329-960-1	Sequence 1, Appli
C 37	32.8	2.7	1298	3	US-08-948-705-3	Sequence 3, Appli
C 38	32.8	2.7	1298	4	US-09-510-543-3	Sequence 3, Appli
C 39	32.6	2.7	245	4	US-09-513-999C-10942	Sequence 10942, A
C 40	32.6	2.7	530	4	US-09-270-767-30636	Sequence 30636, A
C 41	32.6	2.7	911	1	US-08-393-985-5	Sequence 5, Appli
C 42	32.6	2.7	929	4	US-09-270-767-14462	Sequence 14462, A
C 43	32.4	2.7	685	4	US-09-620-312D-806	Sequence 806, App
C 44	32.4	2.7	3635	2	US-08-553-436A-5	Sequence 5, Appli
C 45	32.2	2.7	553	4	US-09-621-976-15491	Sequence 15491, A

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMUN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 6.7%; Score 80.4; DB 1; Length 7218;

Db 165 YSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 106
QY 452 TCTGGCCTCACAGTTGGCCTTCACTGCCTTTTCTCATATTTCTTAACTCGCAAAAGT 511
Db 105 YSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 46
QY 512 TCACTCCTTTGATAGTCAGTTCTTTGCTTCTCTCACT 549
Db 45 YAYAYTTYGYTYAYAYTYAYGYTYAYAYTTYGYT 8

RESULT 4
US-09-244-796-17/c
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

Query Match 3.5%; Score 41.6; DB 3; Length 289;
Best Local Similarity 10.4%; Pred. No. 0.004;
Matches 29; Conservative 108; Mismatches 141; Indels 0; Gaps 0;
QY 272 TTCTGTTCGCTTCTTTTCTCGAATCAGCAACCCAAATCAACAGATACAAATTTCAAGTC 331
Db 285 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 226
QY 332 AGTCCCTTCTTCAACCCCTTGCATCGTTTACTTGTGCACTGCACTGCTAGTCCG 391
Db 225 YCYTGYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 166
QY 392 CTATCTCTATTGTCTGAGTAGGTGCTCTACTTACAGTCTCTACTTTCTCCCTCA 451
Db 165 YSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 106
QY 452 TCTGGCCTCACAGTTGGCCTTCACTGCCTTTTCTCATATTTCTTAACTCGCAAAAGT 511
Db 105 YSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 46
QY 512 TCACTCCTTTGATAGTCAGTTCTTTGCTTCTCTCACT 549
Db 45 YAYAYTTYGYTYAYAYTYAYGYTYAYAYTTYGYT 8

RESULT 5
US-09-621-976-8976/c
; Sequence 8976, Application US/09621976
; Patent No. 6639063

; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8976
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match 3.1%; Score 37; DB 4; Length 399;
Best Local Similarity 12.5%; Pred. No. 0.14;
Matches 46; Conservative 162; Mismatches 157; Indels 2; Gaps 1;
QY 684 AATGCTCTTCAGAAAGTTTTCACGAAGCATACATCCTCAGCAGTCACGGACTTGGCCAT 743
Db 365 AATCAGCCAAAGKWWGAMKWRPTGGWKYVYWMYTSRGSYRYKTSAMNGRAKMKRKT 306
QY 744 TTACCACTCTCTAGTTGCGAGTTGTAGTTCTCATAGACTTTTTCAGAGTGAGTG 803
Db 305 KXWYWMKGGKGGSTYMAIRSRGSTRWSYRRAMWRGSKWGGSYVRYMAGYRSSRW 246
QY 804 GGAACCTTTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 863
Db 245 SWYSAMWRKRWKTCWGRS--SWGSRSTGYAYMYKKSWCTSRKMYKKRKRKRRKCT 188
QY 864 TTTAGCTCGGCAGCTATTTCTCTGCAAGTCTACACTCTGCTGCTGCTGCTGCTGCTGCT 923
Db 187 STKTCYRGSTYKWKAYTKRKRWTWYTYYSYMSKKTWNTKATYATYTKRWKMT 128
QY 924 CGAGTCATCTCTGTTCTTCCATTAACAGCTGTGGAGTTCCTCATATAGTTCAGT 983
Db 127 TKWTCTMCWKCTTYMAGTMMYRYYVYAKRAKSKRCTWSTTCYCMYMAKKWSY 68
QY 984 TCGGCAGTGATAGTTTTCATGATAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1043
Db 67 WMSMSMKWGSMMWKWTYYYYYMMKWSKMTYWSMNSCYARKKRWRTYAKTYTWTMTW 8
QY 1044 AGCTATC 1050
Db 7 GKRTYKY 1

RESULT 6
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match 3.1%; Score 36.8; DB 4; Length 832;

Best Local Similarity 11.8%; Pred. No. 0.24;									
Matches 41; Conservative 149; Mismatches 156; Indels 0; Gaps 0;									
Qy	830	ACTACAAACTCGGGAAGTGCATATGTTTGTGACTTTAGCCTCGGACGCTATTTCTCGGC	889						
Db	26	WNYWYMKTYWRRKKKAWKYKWTWYVRYVMGTYKKACMCRKTKKKKKKG	85						
Qy	890	AAGTCTACACTCCCTGGTCTTGGGATGTATCTTCGAGTATCTCTGTCTTCCAATT	949						
Db	86	YMWYMGWRSSYAMMTTWTGTAYYRSMYMYWRYRCWKKAYRKTTCTSYSSKGTWVK	145						
Qy	950	CCATAACAGCTGGGATTGCTATAGTTGCAGTTGCGGAGTGTATGTTTCCATGATA	1009						
Db	146	RWKAATWTKYKYWAATYVWMMCWTKWRASWYCWGWGAKRWSWTKRSRVSASA	205						
Qy	1010	GAATGGAGCATCAAAATCTCTCCATATATTTAGCTATCTGCGGCTCTCTTCATCG	1069						
Db	206	RSARCCYSCSWGAMSWKMYMRWRWGWATGAGKAWRASCMWRKYAGKSKTSYKSMW	265						
Qy	1070	TCTATCAGCACTACTCGACGAAAGAGTTGAATACTAGCCACACAGTCTGTAGAG	1129						
Db	266	MCWTRSWKYCTKARWTGYCYRKGMMGKGRWYASKYKMYNMKSWWMCWARMYRSTGR	325						
Qy	1130	ATCTTCACTACCTGTTGAGGAAGTGCACAAACATACAAAGTGT	1175						
Db	326	ASMWRRHYITMMKWKYAWARAARWAWWAWRACAAATAT	371						
RESULT 7									
US-09-806-708B-22									
; Sequence 22, Application US/09806708B									
; Patent No. 6784342									
; GENERAL INFORMATION:									
; APPLICANT: The University of British Columbia									
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants									
; FILE REFERENCE: 4810-58741									
; CURRENT APPLICATION NUMBER: US/09/806,708B									
; CURRENT FILING DATE: 2001-04-03									
; PRIOR APPLICATION NUMBER: US 60/147,133									
; PRIOR FILING DATE: 1999-08-04									
; NUMBER OF SEQ ID NOS: 23									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 22									
; LENGTH: 1141									
; TYPE: DNA									
; ORGANISM: Artificial sequence									
; FEATURE:									
; NAME/KEY: promoter									
; LOCATION: (1)..(1141)									
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters									
US-09-806-708B-22									
Query Match 3.0%; Score 35.8; DB 4; Length 1141;									
Best Local Similarity 11.5%; Pred. No. 0.6;									
Matches 84; Conservative 242; Mismatches 395; Indels 7; Gaps 1;									
Qy	399	TTATTGTCTGAGTAGGGTCTCTACTTACAGTCTCTACTTCTCCCTCATCTTGGC	458						
Db	309	WKSRTTKVRTSCWANCTRAGDANKHKWKKWSAAMVYNNNNNNNTYKKARHARDW	368						
Qy	459	CTCAGTTGGCCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACTCC	518						
Db	369	VWHSAAKWHANAHAHSYRKWTBYKRTVMNNNGTTWKKWNAWYKMDMDWGTNNN	428						
Qy	519	TTTGATAGTCAAGTTCTTCTCTCACTGTATCTCTCTCTCTCTCTCTCTCTCTCT	578						
Db	429	NNGRTYGVTKNKKMWYKWKANNCKWRADHKTCTHNNTTWKKMTYNNNCYVKSMT	488						
Qy	579	TGATTCAAGAACTCAACTAATGTATCTAGAGTACAGTATGTGTCGGTTCATATGAC	638						
Db	489	NGKSHRBAAYVTYMMWRRYAHANNNNNDWYKACTWYKBYVCSKWNWNYAAWTKSS	548						
Qy	639	CATCGGTCTCCGCTGGGATTGGACTGTTACTATCTCTGATACAAATGCTCTTCAGAA	698						
Query Match 2.9%; Score 35.2; DB 3; Length 248;									
Best Local Similarity 19.2%; Pred. No. 0.38; Indels 0; Gaps 0;									
Matches 40; Conservative 80; Mismatches 88; Indels 0; Gaps 0;									
Qy	241	CTTCTTCAACTCATTTGGCTTCCCTGTACTGGTCTGTCTCGCTTCTTTCTCGAATCAGG	300						
Db	212	CYTYGYTYTCYCYAYGYTYTYGYTYTYTYCYAYGYCYTYGYTYCYAYGY	153						
Qy	301	CAACCCAAATCAACAGATACAAATTCAGTCAGTCCCTCTCTTACCACCTTGCATCG	360						
Db	152	AYCYGYTYTYTYCYCYAYGYCYAYGYCYTYTYCYTYTYCYTYCYAYGYAYTY	93						

THIS PAGE BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:45:29 ; Search time 3801.29 Seconds
(without alignments)
10228.408 Million cell updates/sec

Title: US-09-913-767-10
Perfect score: 1067
Sequence: 1 agacaagaatggtgaaggct.....gaccataagtgaagttaa 1067

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hc:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gse1:
9: gb_gse2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	440	41.2	1230	CNS0A99D	BX821376 Arabidops
2	381	35.7	989	CL479645	CL479645 SAIL 30 B
3	342.8	32.1	1059	BX841065	BX841065 BX841065
C 4	333.2	31.2	827	BZ515292	BZ515292 BOMQL20TR
5	306	28.7	680	BG439880	BG439880 GA_Ea000
C 6	293.2	27.5	688	BH589732	BH589732 B0GES04TR
7	284.4	26.7	971	CK272858	CK272858 EST718936
8	268.4	25.2	986	CK243751	CK243751 EST727388
9	265.8	24.9	985	CK257466	CK257466 EST741103
10	265.2	24.9	688	AJ807511	AJ807511 AJ807511
11	264.8	24.8	1013	CK251896	CK251896 EST735533
12	262.8	24.6	998	CK247918	CK247918 EST731555
13	261.2	24.5	907	CK257887	CK257887 EST741524
14	246.6	23.1	988	CK248243	CK248243 EST731880
15	246	23.1	979	CK252005	CK252005 EST735642
16	242.4	22.7	722	BM436674	BM436674 VVA008B03
17	238.4	22.3	645	AU238506	AU238506 AU238506
18	236.6	22.2	703	CB006438	CB006438 VVC033611
19	236.4	22.2	930	CK253429	CK253429 EST737066
20	235.6	22.1	942	CK249994	CK249994 EST733631
21	234.8	22.0	929	CK249872	CK249872 EST733509
C 22	234.4	22.0	868	CK245030	CK245030 EST728667
C 23	234.4	22.0	902	CK262518	CK262518 EST708596
24	231.6	21.7	908	CK247701	CK247701 EST731338

25	227.4	21.3	932	7	CK256828	CK256828 EST740465
26	227	21.3	665	6	CD714562	CD714562 VVB202D06
27	225.8	21.2	668	6	CB008408	CB008408 VVC051C02
C 28	225.4	21.1	527	8	AQ958362	AQ958362 LERAX12TR
29	222.2	20.8	928	7	CK245031	CK245031 EST728668
30	222	20.8	948	7	CK257752	CK257752 EST741389
31	221.2	20.7	784	8	BH601144	BH601144 BOHJF33TF
32	219.8	20.6	829	7	CK262517	CK262517 EST708595
33	219.2	20.5	899	7	CK256093	CK256093 EST739730
34	218.6	20.5	894	7	CK247243	CK247243 EST730880
C 35	217.8	20.4	665	5	BQ799268	BQ799268 EST 1437
36	217	20.3	882	7	CK258020	CK258020 EST741657
37	217	20.3	889	7	CK251126	CK251126 EST734763
38	216	20.2	880	7	CK257437	CK257437 EST741074
39	215.2	20.2	616	5	BQ873555	BQ873555 QGI1N19.Y
40	215.2	20.2	701	5	BQ874152	BQ874152 QGI1A120.Y
41	214.2	20.1	892	7	CK246384	CK246384 EST730021
42	213.8	20.0	732	5	BQ988613	BQ988613 QGI15G06
C 43	212	19.9	628	5	BQ798564	BQ798564 EST 733 G
C 44	209.8	19.7	687	5	BQ798954	BQ798954 EST 1123
45	208.2	19.5	257	8	BH812211	BH812211 SALK_0614

ALIGNMENTS

RESULT 1
CNS0A99D
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTSL372D07 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).
ACCESSION BX821376
VERSION BX821376.1 GI:42468600
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1230)
AUTHORS Castellani, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1230)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castellani V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full-length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

Location/Qualifiers
1. 1230
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"


```
Db 925 TTTTCGCCGTCGTTGTTTCCGGTTGTAGTTTTCAGTTAGTGAAGGTCTCTCTACTTC 984
Qy 937 TCTCCCTTTGGGGCTTTGCTCTTACTTTTATGATGAGATAAAGTCG 984
Db 985 TTTACATTTGGGGTTTGTCTCATACATCTAAGGTGTGTTTAAATCGG 1032

RESULT 4
BZ515292/c
LOCUS BZ515292 BO_2.3 KB Brassica oleracea genomic clone BOMQL20,
DEFINITION genomic survey_sequence.
ACCESSION BZ515292
VERSION BZ515292.1 GI:27043827
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 827)
Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other_GSSs: BOMQL20TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES             source
    Location/Qualifiers
        1..827
            /organism="Brassica oleracea"
            /mol_type="genomic DNA"
            /strain="TO1000DH3"
            /db_xref="taxon:3712"
            /clone="BOMQL20"
            /clone_libs="BO_2.3 KB"
            /note="Vector: pHC51, Site 1: BstXI; 2-3 kb sheared
            genomic DNA inserted into pHC51 using BstXI linkers"

ORIGIN
Query Match 31.2%; Score 333.2; DB 8; Length 827;
Best Local Similarity 67.1%; Pred. No. 1.9e-82;
Matches 546; Conservative 0; Mismatches 239; Indels 30; Gaps 4;

Qy 1 AGACAAGAATGGTGAAGGCTTTGTGATCATAAACTGCATAATTTCTAGCCATAGGAAC 60
Db 791 AGAAGAAGTGAAGACTGGTCTTGTAATCATAAACTGTATAATCTCGCTATTGGAACT 732
Qy 61 GTGGAGTCCCTTTGATTATGGTCTCTACTTCAACAATGGCGGTAAGAAGTTTGGTCT 120
Db 731 GCGGAGGCCCTCTAATACCGGTCTCTACTTTCAGAAACGGTGGCAACGGATCTGGTCT 672
Qy 121 CTACGTTTCTTGAACACTGCAGCTTTCCTGTATCTTCATTCCTCTGCTCTCTCTACA 180
Db 671 CAAGCTTCTCCAAACCGAGGCTGCCATCATCCCTCTCCCTCTCTCTCTCTCTTC- 613
Qy 181 TTACCCGCGCAGAGCAACAATGTGGGTGATGATGACAAAGTTTCTTTCTTATCAAAACGC 240
Db 612 CTCGCCGCTAAGAAGCAACAACG-----ACTCCATTTTCTCTATAAAACCTC 564
Qy 241 GTCCTCTTATCGCGCTGTTATTGTAGGCATCTCTCAGGCTTGTGATAACTACTTGTATG 300
Db 563 CTCGTGTTCTTGCGCTTCTATTGTGATGTGCTCTGCTACGTGGCGGTGACAAATACCTACG 504
Qy 301 CATATGGTATAGCTTATCTCCAGTTTCTACAGCTGCTCTTATCATTTGCTTTCACAGTTAG 360
Db 503 CTTAGGGGTAGCTTATACCTCCAGTTTCCACTTCTCTTCTTGTGTCATCTCCTCCCACTGG 444
```

```
Qy 361 CTTTTATAGCTATCTTCTCATTTCTTCAAGTTAACATAAAGTTCACTCTCTTTTACCATCA 420
Db 443 CCTTCATCTGCTTTCTTGGCCTTTTTCATGGTGAAGCAAAAAGTTTCAGCCATTTATTATAA 384
Qy 421 ATGCTGTGTGTTGTTGACTGTTGGTGTCTGCGGTTTTTGGGAATGTCATACCGAAACTGATA 480
Db 383 ACGCCGTCGTTTGTCTTACTCTCGGTGCCGGAATCCTTGGCCCTTCACACCGACGGTGACA 324
Qy 481 AGCAGTTTATAGACTCACAAAGCAGTACATAAATCTGTTTCTTGTACTGTAGCAGCAG 540
Db 323 AGCTTGCCAAAGACACACAAAGAGTATATCGTTGGTTCTCTCATGACCCCTTGTGCGG 264
Qy 541 CTGTTATGTATGCTTTTCATCTTGGCCATTAGTGGAACCTTGCTTACCAGAAAGCTAAGCAAA 600
Db 263 CTATTCTCTACGGGTTTGTATTGCCGCTTGTGGAGCTCACTTACAAGAAAGCTCGTCAAC 204
Qy 601 CCATGAGCTATACCCCTTGTGCTCGAGTCCAGTTCCAGTTGATTTTGTCTCTCTTGTCTATTG 660
Db 203 GAATCAGCTACACGCTTGTCTCGAGATGACGATGCTTGTGCTTGTGCTGCTGCCACTTGCT 144
Qy 661 TCAGCCTCATCGGTATGTTTCATCGCTGGTGATTTCAGCAGGCGCTTACCAAAAAGAACAA 720
Db 143 TCTGCTCGTTGGATGCTAGCTGATGGCGATTTCAAGGAG-----TTAGCAA 96
Qy 721 GAGAGTTCAAGCTTGGAGAG-----GCATTGTTCTATGTGGTGGCTGTGTTTTCAGCCA 774
Db 95 GAGATTTTAAAGCTTGGAGGCTCTACTACGTATTACTATGTGTTGTTGTGTTTTCAGGCCA 36
Qy 775 TCATATGGCAAGGCTTCTTCTTGGGAGCCATTGG 808
Db 35 TCAITCGGCAAGGCTTTTCTTGGGATCTATTGG 2
```

```
RESULT 5
BG439880
LOCUS BG439880
DEFINITION GA_Ea0005C03f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboresum cDNA clone GA_Ea0005C03f, mRNA sequence.
ACCESSION BG439880
VERSION BG439880.1 GI:13349537
KEYWORDS EST.
SOURCE Gossypium arboreum
ORGANISM Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 680)
Wang,R.A., Friesch,D., Yu,Y., Main,D., Rambo,T., Simmons,J.,
Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rtwing@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
High quality sequence stop: 621.
FEATURES             source
    Location/Qualifiers
        1..680
            /organism="Gossypium arboreum"
            /mol_type="mRNA"
            /strain="AKA"
            /cultivar="8400"
            /db_xref="taxon:29729"
            /clone="GA_Ea0005C03f"
            /tissue_type="Fibers isolated from bolls harvested 7-10
            dpa"
            /lab_host="E. coli"
```


CK272858	971 bp	mrna	linear	EST 03-AUG-2004
LOCUS	EST7119936	potato abiotic stress	cdna library	Solanum tuberosum cdna
DEFINITION	clone POAD778 5' end, mRNA sequence.			
ACCESSION	CK272858			
VERSION	CK272858.1	GI:39829836		
KEYWORDS	EST.			
SOURCE	Solanum tuberosum (potato)			
ORGANISM	Solanum tuberosum			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanaceae; Solanaceae; Solanum.			
AUTHORS	1 (bases 1 to 971)			
TITLE	Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.			
JOURNAL	Generation of ESTs from abiotic stressed potato tissue			
COMMENT	Unpublished (2003)			
	Other ESTs: EST718937			
	Contact: Robin Buell			
	The Institute for Genomic Research			
	9712 Medical Center Dr, Rockville, MD 20850, USA			
	Email: potato-array@tigr.org			
	Clones can be requested from the University of Arizona Genomics			
	Institute via http://genome.arizona.edu/orders/			
	Seq primer: ATT TAG GTG ACA CTA TAG.			

FEATURES

```

/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clones="POAD778"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="potato abiotic stress cDNA library"
/notes="vector: pCMWsport6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

```

ORIGIN

	Query Match	26.7%	Score 284.4;	DB 7;	Length 971;
	Best Local Similarity	58.2%;	Pred. No. 1.2e-68;		
	Matches 559; Conservative	0;	Mismatches 391;	Indels 10;	Gaps 3;
Qy	9	ATGGTGAAGGCTTTGTGATCATAACTGCATTAATTCTTAGCCATAGAAAACCTCGAGGT	68		
Dd	20	ATGAGGAGATTCCCTCTTGATTATAAATTGTTTACTATCCGTGCTGCTTTCTGTGGC	79		
Qy	69	CTTTGATTATGCGTCTCTACTCAACAATGGCGGTAAAGGATTTGGTTCCTCAGTTT	128		
Dd	80	CCTTAATGATGCGTTTATATTTTGTGAAGGTGTTTTAAGATTATGGTTTAATAGTTGG	139		
Qy	129	CTTGAAACTGGAGGCTTCCCTGTTATCTTCATTCCTGCTCTCTCTTACATTACC	188		

Db	140	TTACAAACCGGTGATGCCCACTCACAAATTATACCAATTAGTCACTTATATTTCTATACGA	199
Qy	189	CGCAGAAGCAACAATGTGGGTGATAGTACAAGTTTCTTTCTTATCAAAACCGCGTCTTCTT	248
Db	200	CGAAAAACCAA-----GGGCTCTGATACCAAGTTTTATTTTATACACCTCGTATTTTC	253
Qy	249	ATCGCGCTGTTATTTGTAGGCATTTCTCAGGGTTTGTAACTACTTGTATGCATATGGT	308
Db	254	ATTGCATCGTTCAATATTATGGCGTTTTACGGGTGCTGATTTCTTACCTTTATTTCATGGGGC	313
Qy	309	ATAGCTTATCTCCAGTTTCTACAGCTGCTCTTATCAATGCTTTCACAGTTTAGCTTTTATA	368
Db	314	GGGTGAAACTCCCCGTTCAAAGTCTCTCATCGAGCTCAACTGTCCTTCACG	373
Qy	369	GCTATCTTCTCATTTTCATGGTTAAACATAAAGTTCACCTCTTTTACCATCAATGCTGTT	428
Db	374	CGCATAGGGTCTTACTTTTCATAGCAAAAGATAAAATTTACATCATATTCGATTAACGCGGTG	433
Qy	429	GTGTTGTTGACTGTGGTGTCTGGGTTTTGGGAATGCATACCGAAACTCATAGCCAGTT	488
Db	434	GTTTTATGTACACTTGGCGCGGTTTTATGGGTATGCGATCGAACGGTATCGACCGGAG	493
Qy	489	CATGAGACTCACAGCAGTACATAAAGTGTGTTCTTGATTACTGTAGCAGCAGCTGTTATG	548
Db	494	GGTGTCAAGTAAAGATATGTTATTTGGTTTTATTATGACACTTTTGGCTGCAGCTTTG	553
Qy	549	TATGCTTTTCATCTTGCCATTTAGTGGAACTTGCTTACCAGAAAGCTAAGCAACCATGACC	608
Db	554	TATGGACTTATTTTACCTGTTATTTGAGTTGATATATTTGAAGGCAAAACAAGCTATTACT	613
Qy	609	TATACCCCTTGTGCTCGAGTTCCAGTTGATTTTGTGCTCTCTGCTTCTATTTGTCAGCGTC	668
Db	614	GCAACATTGGTGTGGAGATTCAAATGGTCATGGCTTTTGGCTACTGCTTTTTCGACA	673
Qy	669	ATCGGTATGTTCACTGCTGGTGATTTTCAAGCAGGCCCTTACCAAAAGAGCAAGAGAGTTC	728
Db	674	ATTGGAAAG--ATTGTCTACAAAGATATTTTCAAGCAATATCAAGGGAAGCAAAACATTTT	730
Qy	729	AAGCTTCGAGAGCATTTGTTCTATGTGTGGCTGTGTTTTCAGCCCATCATATGGCAAGGC	788
Db	731	AACCTTGGAGAAGTAGATATACATAGTTGTAAATGGAATGCCAATATTTGGCAATGT	790
Qy	789	TTCTTCTTGGAGCCCATTTGGATTAATCTTCCACATCGTCTCTCGTCTCGGATATTATG	848
Db	791	TTCTTTTGGTGTCTGTGGAGTTATTACCTCTCTTCTTTTAATGTCCTGAGTTATG	850
Qy	849	ATATCAGTGTCTTTTGGCAATTACAGAGTTTTAGCTGTTATATCTTACCATGAAAGTTT	908
Db	851	ATTGCAGTTTATCTGCGCAATTTACTGAAGTATTAGGAGTAATTTTCTTGTATGAAAAATCT	910
Qy	909	CAAGCTCAGAAGGAGCTTTCTTGTGCTCTCCCTTTGGGGCTTGTCTCTTACTTTTAT	968
Db	911	C-AGCTGAAAGGAGCTTTCACATTTTCTTCTCTTTGTTGGGTTGTTTCTTATTTTT	969
RESULT 8			
CK243751			
LOCUS			
DEFINITION	EST727388 potato callus cDNA library, mRNA linear EST 30-JUL-2004		
	Solanum tuberosum cDNA clone POCAS40 5' end, mRNA sequence.		
ACCESSION	CK243751		
VERSION	CK243751.1		
KEYWORDS	GI:39788661		
SOURCE	EST.		
ORGANISM	Solanum tuberosum (potato)		
	Solanum tuberosum		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	asterids; lamids; Solanales; Solanaceae; Solanum.		
	1 (bases 1 to 986)		
REFERENCE	Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.		
AUTHORS	Generation of ESTs from potato callus tissue		
TITLE	Unpublished (2003)		
JOURNAL	Contact: Robin Buell		
COMMENT			

The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source
1. .986
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="FOCA540"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 25.2%; Score 268.4; DB 7; Length 986;
Best Local Similarity 57.8%; Pred. No. 3.8e-64;
Matches 552; Conservative 0; Mismatches 406; Indels 10; Gaps 3;
QY 25 TGATCAATAACTCATAAATCTAGCCATAGGAACTGTGGAGTCTCTTTGATTATCGGTC 84
DB 27 TAATTTTCACATAATATGCTTACAAATGCGAATTTGGATCCAGCGTATTACAACTGTTGGTT 86
QY 85 TCTACTTCAACAATGCGGTAAGGATTGGTTCTCTACGTTTCTTGAACATGCAAGGCT 144
DB 87 TCTATTTCATTCATGCGGCGCAACGAATTTGGATACCCAGCGTATTACAACTGTTGGTT 146
QY 145 TTCTCTTATCTTCT 204
DB 147 GTCCAAATATTCATCCCTCTAGCCATAGCCCTATTTCCAACGCCGGAATTTCAAGGAC 206
QY 205 TGGGTGATAGTACAAGTTCTTTCTTATCAACCGCGTCTCTTATCGCGCTGTTATTG 264
DB 207 CCGAGCTAAA-----TGGTCTTATCACGGCAAGTTCATCGATCTGCGCGG 260
QY 265 TAGGCATCTCTCAGGGTTGATACTCTGATGATGATGATGATGATGATGATGATGATGATGAT 324
DB 261 TCGGAATCATCGTCGCTCGATCGTCTTAACTCATGGGACCGCGCAATACCGG 320
QY 325 TTCTACAGCTGCTCTTATCT 384
DB 321 TTTCAACTTCGACACTAAATCAACGCGACTCACTTCGCTTCTACTGCGCTTTTGTGTCG 380
QY 385 TCATGTTTAACTAATAAGTTTCACTCTCTTTTACCAATCAATGCTGTTGTTGTTGTTGTTGTTG 444
DB 381 TTATAGTTAAACAGAATTTGACAGCGTATTCGACGAATTCGCTGTTTGTCTTATCGCGC 440
QY 445 GTGCTCGGTTTGGGAATGCATACCGAACTGATAAGCCAGTTTCATGAGACTCAACAAC 504
DB 441 GAGCTCGCACTTTAGCTCTCGGCGCAACGGTGACCGCGCGCGCGCGAGTCCAGAAAG 500
QY 505 AGTACATACTGTTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 564
DB 501 ATTATATGTTGGGGTTTGTATGACGTTTATCGTGGTGTGTTATATGATGATGATGATGATGAT 560
QY 565 CATTAGTGAACCTGCTTACCAAGAGCTAAGCAACCATGAGCTATACCTTGTGCTCG 624
DB 561 CGTTAATGAGTTGATTTATATGAGCGCAAGCACTGTTTACTTACATACAGTATG 620
QY 625 AGTTCAGTTGATTTTGTCT 684
DB 621 AGATTACAGATGTTTGGGCAATTTCTGCTACTGTTTTTGGCACTATTTGGAATGATATCA 680
QY 685 CTGGTGAATTCAGAGCGCTTACCAAGAGCAAGCAAGAGTTCAAGCTTTGAGAGGCAAT 744
DB 681 ACAAGGATTTTCCAGGGGATTTCCAGGGAAGCAAGTCAATATGCAATTTGGAGATCTA 737

QY 745 TGTCTCTATGCTGGTGTGTTTTTTCAGCCATCATATGGCAAGGCTTCTTCTTGGGAGCCA 804
DB 738 AGTATTATATGCTACTAGTATGCTGTCATTTTGGCAATTCGGGCTCTCGGGCTTAG 797
QY 805 TTGGATTATCTCTTCCACATGCTCTCTCGTCTCGGTTATATGATATCAGTGTCTTTTGC 864
DB 798 TTGGAGTTATTTTTTTTATTTCTCTTACTCTCTGGAATATATGCGGCTTTTTTACTTTC 857
QY 865 CAATTACAGAGGTTTTTAGCTGTATATCTACCATGAAAGTTTCAAGCTGAGAGGGAC 924
DB 858 CTGTTACCGAAGTTTGGCTGTAATTTGTTCCAGAAAATTTCAAGCCGAAAAGGAG 917
QY 925 TTTCTCTTCT 983
DB 918 TAGCTATTTTCTAGCTCTATGCGGGAATTTGTTTCATATTTTATGTTGAATATGAACA 977
QY 984 GCGGAGGA 991
DB 978 GCGGAAGA 985

RESULT 9

CK257466 985 bp mRNA linear EST 30-JUL-2004
LOCUS EST741103 potato callus cDNA library, normalized and full-length
DEFINITION Solanum tuberosum cDNA clone POCDS57 5' end, mRNA sequence.
ACCESSION CK257466
VERSION CK257466.1 GI:39814446
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 985)
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
TITLE Generation of ESTs from potato callus tissue
JOURNAL Unpublished (2003)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source
1. .985
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCD557"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 24.9%; Score 265.8; DB 7; Length 985;
Best Local Similarity 56.8%; Pred. No. 2.1e-63;
Matches 533; Conservative 0; Mismatches 397; Indels 9; Gaps 2;
QY 25 TGATCAATAACTGCAATAATCTTAGCCATAGGAACTGTGGAGTCTCTTTGATTATCGGTC 84
DB 56 TAAATTTCAACATAATATGCTTTCAATCGCAATTTGGTGTCTCTTAACTCTCTCGAC 115
QY 85 TCTACTTCAACAAATGCGGTAAGGATTGTTCTCTACGTTTCTTGAACATGAGGCT 144
DB 116 TCTATTTTCAATTCATGCGGCCCAACGAATTTGGATACCCAGCGTATTACAACTGTTGTT 175

```
QY 145 TTCCCTGTTATCTTCAATTCCTCTGCTCTTCTTACATTACCGCGCGCAGAACACAAATG 204
DB 176 GTCCCAATTATCTCATCCCTCTAGCCATAGCCTATTTCCACGCCGGAATTTCAAGAC 235
QY 205 TGGGTGATAGTACAAGTTTCTTTCTTATCAACCCGCTCTTCTTATCGCCGCTGTATTG 264
DB 236 CCGAGGCTAAAA-----TCGTCCTTTATCACAGCGCAAGAGTTTCATCGCATCTCGCGCG 289
QY 265 TAGGCATTTCTCAGGGTTTGAATACTACTTGTATGCAATATGGTATAGCTATCTTCCAG 324
DB 290 TCGGAATCATCGTCGGTCTCGATGTTTACTTAAACTCATGGGAGCCCGCAAAATTACCCG 349
QY 325 TTTCACAGCTGCTCTTATCATTTGCTTCACAGTTAGCTTTTATATAGCTATCTTCTCATTTCT 384
DB 350 TTTCACCTTCGACATAATACAGCGACTCAACTTGGCTTCACTCGGCTTTTGTGTGTC 409
QY 385 TCATGGTTAAACATAAAGTTCACTCTCTTTTACCATCAATGCTGTGTGTGTGTTGACTGTTG 444
DB 410 TTATAGTTAAACAGAAATTGACAGGTATTCGACGAATTCGGTCTGTTTGTCTTATCGCGC 469
QY 445 GTGCTGCCGTTTGGGAATGCATACCGAACTGATAAGCCAGTTTCATGAGACTCACAAAGC 504
DB 470 GAGCTGCCACTTTAGCTCTCCGGCGGAACGGTGACCGCGCGCGCGAGTCCAGAAAG 529
QY 505 AGTACATAAATCGTTTCTTGATTACTGTAGCAGCAGCTGTTATGATGCTTTTCATCTTGC 564
DB 530 ATTATATGTTGGGTTTGTGATGACGTTATTCGGTTCGGGTGTTATATGATGATGTTGTC 589
QY 565 CATTAGTGGAACTTGCTTTACCAGAAAGTCAAGCAACCATGAGCTATACCCCTTGTGCTCG 624
DB 590 CGTTAAATTTGAGTTGATTTATATGAAGCGAAGCAAGCTGTACTTTACACTACAGTATTGG 649
QY 625 AGTTCAGTTGATTTTGTCTCTCTGCTTCTATTTGTCAGCGTCAATCGGTATGTTCAFCG 684
DB 650 AGATTACAGATGGTTTGGGCAATTTCTGCTACTGTTTGTGCACTATTGGAAATGGTTATCA 709
QY 685 CTGTGTGATTTCAAGCAGCGCTTACCAGAAAGCAAGAGAGTTCAAGCTTCGAGAGGCAT 744
DB 710 ACAGAGATTTC---CAGCGAATTCNAGGAGCAAGTCAATATGCAATTCGAGAACTA 766
QY 745 TGTTCTATGTGGTGCTGTGTTTCAGCCATCATATGGCAAGGCTTCTTCTTGGGAGCCA 804
DB 767 AGTATTATATGCTACTAGTATGGTGTGCCATTTATTTGGCAATTCGCGCTCTGGGGTTAG 826
QY 805 TTGGATTAATCTTCCACATCGTCTCTGCTCTCGGTATTTATGATATCAGTGTCTTTGTC 864
DB 827 TTGGAGTTATTTTATTTCTTCTCTTACTCTCTGGAATTTATAGGCGCTTTTCTTACTTC 886
QY 865 CAATTACAGAGGTTTATGCTGTATATTTCTACCATGAAAGTTTCAAGCTCAGAGAGGAC 924
DB 887 CTGTTACCGAGTTTGGCTGTAAATTTGTTCCAGAAATAATTTCAAGCCGAAAGAGAG 946
QY 925 TTTCCTGTGCTCTCTCCCTTTGGGGCTTTGCTCTTACT 963
DB 947 TAGCTATTTTCTAGCTCTATGGGATTTGTTTCATATT 985
```

RESULT 10

AJ807511

LOCUS

DEFINITION AJ807511 Antirrhinum majus whole plant Antirrhinum majus cDNA clone 018.6.06.p14, mRNA sequence.

ACCESSION AJ807511

VERSION AJ807511.1

KEYWORDS

SOURCE

ORGANISM

Antirrhinum majus (snapdragon)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Plantaginaceae; Antirrhineae;
Antirrhinum.

REFERENCE

1 (bases 1 to 688)

AUTHORS

TITLE

JOURNAL

COMMENT

Zachgo, S., Stueber, K., Saedler, H., Sommer, H. and Schwarz-Sommer, Z.
Antirrhinum EST collection
Unpublished (2003)
Contact: Schwarz-Sommer Z
Molekulare Pflanzen-genetik
MPI fuer Zuechtungs-forschung
Carl-von-Linne Weg 10, D-50829, Germany.
Location/Qualifiers

FEATURES

source

1..688

/organism="Antirrhinum majus"

/mol_type="mRNA"

/db_xref="taxon:4151"

/clone="018.6.06.p14"

/tissue_type="whole plant"

/clone_lib="Antirrhinum majus whole plant"

ORIGIN

Query Match 24.9%; Score 265.2; DB 1; Length 688;

Best Local Similarity 64.0%; Pred. No. 2.7e-63;

Matches 432; Conservative 0; Mismatches 239; Indels 4; Gaps 2;

QY 330 ACAGCTGCTCTTATCATGTTCACAGTTAGCTTTTATAGCTATCTTCTCATTTCTCATG 389

DB 11 ACCTCGTCACTGATTAATTCGACGCACTTGGCGTTTACTGCGGGTTTGCCTTATTTCTT 70

QY 390 GTTAACATAAAGTTCACCTCTTTTACCATCAATGCTGTGTGTGTGTTGACTGTGTGCT 449

DB 71 GTGAAGCAGAGATTTCACGGCGTTT-TNACAAACGATCGTTTTTGTGTGACGTTGGGAGCG 129

QY 450 GCGGTTTTGGGAATGCATACCGAAACTGATAAGCCAGTTTCATGAGACTCAACAAGCAGTAC 509

DB 130 GTGGTCTAGGGCTGCACACAGCGCGATAGGCCGGAGGAGAGTGAACAAAGAGTAT 189

QY 510 ATAACTGGTTTCTGATTAATCTGTAGCAGAGCTGTTATGATGCTTTCACTTGTGCCAATTA 569

DB 190 TTAGCTGGGTTTTTATGACGCTTGGGCGCGCGCTTATGCGTTTATTTGCGCATTTG 249

QY 570 GTGGAACCTGCTTACCAAGAAAGCTAAGCAAAACCATGAGCTATACCTTGTCTCGAGTTC 629

DB 250 GTGGAGTTGACGTATATGAAGCGGAGGCGGCTTGAGTTACACACTCGTATTGGAAAT 309

QY 630 CAGTTGATTTTGTCTCTCTCTCTCTTCTTATTTGTGAGCGTCTATCGGTATGTTTCATCGCTGGT 689

DB 310 CAGTTGCTCATGTTCTTCTCGCACCTGCTTTTGTACTGTTGGAATGCTCGTCAACAAA 369

QY 690 GATTTCAAGCAGGCTTTACCAAGAAAGCAAGAGAGTTCAAGCTTGGAGAGGCAATGTTTC 749

DB 370 GATTTTC---CAGGCAATTCCAAGAGAAAGCAAGCGTACGAGCTAGGGGAGACAAAATAC 426

QY 750 TATGTGGTGGCTGTGTTTTTCAGCCATCATATGGCAAGCTTCTTCTTGGGAGCAATTTGA 809

DB 427 TATTAGTGGTAGTGTATTAGTGCATAAATTTGGCAGTGTCTTTCTTGGGAGCAATTTGA 486

QY 810 TTAATCTTCTCCACATCGTCTCTCGTCTCGGGTATTATGATATCAGTGTCTTTTGGCCAAAT 869

DB 487 ATCATATTTTACTCTTCGCTCTCTTTTATCGGCGATCGTAATTAATCTTACTTCTTCCAAATC 546

QY 870 ACAGAGGTTTTAGCTGTATTAATCTACCATGAAAGTTTCAAGCTGAGAGAGGCACTTTCT 929

DB 547 ACAGAAATTTGTGGCTGTATTATTTTACCATGAAATAATTTCAAGCAGAAAAAGGATTTTCT 606

QY 930 CTTGCTCTCTCCCTTTGGGCTTGTCTCTTACTTTTATGTTGAGATAAAGTCTGGCGAG 989

DB 607 CTTTCTCTCTCGGTGTGGGATTTTGTCCGTACTTTTATGTTGAGATCAAAACACTTGAAG 666

QY 990 GATAAAGAGGAGAAAT 1004

DB 667 AAGATGAAGGGAAGT 681

RESULT 11

CK251896

LOCUS

CK251896

1013 bp

mRNA

linear

EST 30-JUL-2004

```

DEFINITION EST735533 potato callus cDNA library, normalized and full-length
ACCESSION Solanum tuberosum cDNA clone POCBR03 5' end, mRNA sequence.
VERSION CK251896
KEYWORDS CK251896.1 GI:39805365
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 1013)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
Generation of ESTs from potato callus tissue
Unpublished (2003)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
    Location/Qualifiers
        1..1013
            /organism="Solanum tuberosum"
            /mol_type="mRNA"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone="POCBR03"
            /tissue_type="callus"
            /lab_host="DH10B-Tona"
            /clone_lib="potato callus cDNA library, normalized and
            full-length"
            /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
            supplier: RNA was isolated from Solanum tuberosum var.
            Kennebec callus tissue grown on solid media."

ORIGIN
Query Match 24.8%; Score 264.8; DB 7; Length 1013;
Best Local Similarity 57.2%; Pred. No. 4e-63;
Matches 543; Conservative 0; Mismatches 397; Indels 10; Gaps 3;

QY 25 TGATCATAACTGCATAATCTAGCCATAGGAACATGTGGAGTCTCTTTGATTATCGGTC 84
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
63 TAAATTTTCAACATAATATGCTTTTCAATCGGAATTTGGGTGCTCTTTAATCTCTCGAC 122
QY 85 TCTACTTCAACAAATGCGGTAAAGGATTTGGTTCTCTACGTTTCTTGAACCTGCAGGCT 144
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123 TCTATTTTCAATTCATGGCGCCACGAAATTTGATACCCAGCGATTAACAACCTGTGGTT 182
QY 145 TTCTCTGTTATCTTCATTCCTCTGCTCTCTCTTCAATTAACCGCGGCGAAGCAACAATG 204
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
183 GTCCAAATATTCTATCCCTCTAGCCATAGCCTATTTCACACGCCGAAATTCAGGAC 242
QY 205 TGGGTGATAGTACAAGTTCTTTTATCAACACCGGCTCTTATTCGCGCTGTATTG 264
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
243 CCGAGGCTAAAA-----TCGCTTTATCACGCGCAAGATTCATCGCATCTGCCGCG 296
QY 265 TAGGCAATCTCTCAGGGTTTGATACTACTTGTATGCATATGATAGCTTATCTCCAG 324
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
297 TCGGAATCATCTCGGTCTCGAATGTTTAACTATGCGGACCCCGGAAATTAACCGG 356
QY 325 TTTTACAGCTGCTCTTATCATTTGCTTTCACAGTTAGCTTTTATAGCTATCTTCTCATTTCT 384
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
357 TTTCAACTTCGACATCAATCAACGCGACTCACTTGGTTCACCTGGCTTTTGTGTGTC 416
QY 385 TCATGGTTAAACATAAAGTTCACTCTCTTTTACCAATCAATGCTGTTGTGTTGACTGTTG 444
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
417 TTATAGTTTAAACAGAAATTCAGACGCTATTCGACGAATTCCTGCTGTTTGTCTATCGCG 476
QY 445 GNGCTCGGTTTTGGGAATGCATACCGAACTGATAAGCCAGTTTATGAGACTCAAGC 504
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
477 GAGCTCGCACTTTAGCTCTCCGGGCGAACGGGTGACCGCGCCCGCGGAGTCCACGAAG 536

```

ORIGIN


```
QY 550 ATGCTTTTCATCTTGGCCATTAGTGAACCTGCTTACACGAAAGCTAGACAAACCATGAGCT 609
Db 478 ATGGAATGATGTTGCCGTTAAATTTAGTGTGATTTATATGAAAGCGCAAGCTGTTACTT 537
QY 610 ATACCCCTTGTCGCGAGTTCAGTTGATTTTGGTCTCCTTGTCTTCTATTGTTCAGGCTCA 669
Db 538 ACACTACAGATTGGAGATTCAAGTGGTTTGGGCAATTCCTGCTACTGTTTTTGGCACTA 597
QY 670 TCGGTATGTTTCATCGCTGTTGATTTCAAGCAGGCTTTACCAAGAGCAAGCAAGAGATTCA 729
Db 598 TTGGAATGATTATCAACAAGGATTC---CAGGGATTCCAAAGGAAGCAAGTCATATG 654
QY 730 AGCTTGAGAGGCAATGTTCTATGTGGTGTGCTGTTTTCAGGCATCATATGCAAGGCT 789
Db 655 CAATTGGAGAATCTAAGTATTATATGTTACTAGTATGTTGTCGCAATTTATTTGCAATTCG 714
QY 790 TCTTCTTGGGAGCCATTGGAATTAATCTTCCACATCGTCTCTCGTCTGGGTATTATGA 849
Db 715 CGCTCTGGGGTAGTTGGAGTTATTTTATTTCTCTTCTTCTCTCTCTCTCTCTCTCTCTCT 774
QY 850 TATCAGTGTCTTTTGGCAATTACAGAGTGTATAGTGTATATATTTTCTACCATGAAAGTTTC 909
Db 775 GCGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 834
QY 910 AAGCTGAGAGGAGCTTTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 969
Db 835 AAGCCGAAAGAGTAGCTATTTTCTAGCTCTATGGGGAATTTGTTTCATATTTTATG 894
QY 970 GTGA 973
Db 895 GTGA 998

RESULT 14
CK248243
LOCUS
DEFINITION
EST731880 potato callus cDNA library, normalized and full-length
Solanum tuberosum cDNA clone POCAZ44 5' end, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 988)
Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
Generation of ESTs from potato callus tissue
Unpublished (2003)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
Location/Qualifiers
1..998
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCAZ44"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."
```

ORIGIN

RESULT 15

CK252005

LOCUS

DEFINITION

CK252005 979 bp mRNA linear EST 30-JUL-2004

DEFINITION

EST735642 potato callus cDNA library, normalized and full-length

Solanum tuberosum cDNA clone POCBR86 5' end, mRNA sequence.
CK252005
VERSION CK252005.1 GI:39805578
KEYWORDS EST,
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 979)
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
TITLE Generation of ESTs from potato callus tissue
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST735643
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via <http://genome.arizona.edu/orders/>.
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
source location/Qualifiers
1..979
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCBR86"
/tissue_type="callus"
/lab_host="DH10B-TonA"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 23.1%; Score 246; DB 7; Length 979;
Best Local Similarity 56.4%; Pred No. 8e-58;
Matches 523; Conservative 0; Mismatches 395; Indels 10; Gaps 3;
QY 25 TGATCATAACTGCATAATCTTAGCCATAGGAAACTGTGGAGGTCCTTTGATTATGCGTC 84
Db 62 TAAATTTCAACATAATATATGCTTTCAATCGGCAATTCGGTGGTCTTTAATCTCTCGAC 121
QY 85 TCTACTTCAACNATCGCGGTAAAGGATTTGGTCTCTACGTTCTTGAAACTCGAGGCT 144
Db 122 TCTATTTTCATTCATGCGGCGCAACGAATTTGGATACCCAGCGTATTCACAACTGTGGTT 181
QY 145 TTCCCTGTTTCTTCAATCTCTGCTCTCTTCTTACATTCACCGCGCAGAGCAACATG 204
Db 182 GTCCAAATTAATCTCAATCCCTCTAGCCATAGCCTATTTCCAGCGCGGAAATTCAGGAC 241
QY 205 TGGGTGATAGTACAAGTTCTTTCTTATCAACCGCGTCTTTTATCGCCGCTGTATTG 264
Db 242 CCGAGGCTAAAA-----TGGTCTTTATCACAGCGCAAGAGTTTCATCGCATCTGCGGCG 295
QY 265 TAGGCATTTCTCAGGGTTTGATTAAGTACTTGTATGCAATATGGTATAGCTTATCTTCAG 324
Db 296 TCGGAATCATCGTCGGTCTCGATGGTTACTTAAACTCATGCGGAGCCCGGAAATTCACCG 355
QY 325 TTTCCTACAGCTGCTCTTATCATTTGCTTCACAGTTAGCTTTTATAGCTATCTTCTCATTTCT 384
Db 356 TTTCACCTTCGACTTAATCAACGGCACTCACTTGGCTTCACTGCGCTTTTGTCTGTCG 415
QY 385 TCATGGTTAACAATAAAGTTCACTCTCTTTTACCATCAATGCTGTGTGTGTGTTGACTGTG 444
Db 416 TTATAGTTAAACAGAAATTCAGCGGTATTCGACGAATTCGGTCTGTTTGTCTATTCGCGC 475
QY 445 GTGCTCGGGTTTTGGGAATGCATACCGAAACTGATAAGCAGTTCATGAGACTCACAGC 504
Db 476 GAGCTGCGACTTTAGCTCTCCGGGCGAAACGGTGAACCGCGCGCGCGAGTCCACGAAG 535

QY 505 AGTACATAAATGGTTTCTTGATTACTGTAGCAGCAGCTGTATGTATGTATGTCTTCACTTGC 564
Db 536 ATTATATGTTGGGTTTGTGATGACGTTTATCGGTGGGTGTATATGGATTGATGTTGC 595
QY 565 CATTAGTGAACCTTGCTTACAGAAAGCTAAGCAACCATGAGCTATACCCCTGTGTCTCG 624
Db 596 CGTTAAATTGAGTTGATTATATGAAGCGCAAGCAAGCTGTACTTACACTACAGTATTGG 655
QY 625 AGTTCCAGTTGATTTTCTGCTCTCTTCTCTTCTTCTTGTGACGTCATCGGTATGTTTCATCG 684
Db 656 AGATTCAGATGGTTTTGGGCATTTCTGCTACTGTTTTTGTGCACTATTGGAAATGATTATCA 715
QY 685 CTGGTGATTTCAAGCAGGCTTTACCAAAAGAACAGAGAGTTTCAAGCTTTGGAGAGCAT 744
Db 716 ACAAGGATTTTC---CAGGCGATTCCAAGGGAAGCAAGTCAATATGCAATTTGGAGATCTA 772
QY 745 TGTTCTATGTTGGGTGCTGTTTTCAGCCATCATATGGCAAGGCTTCTCTTGGAGGCCA 804
Db 773 AGTATTATATGGTACTAGTATGGTGTGCCATTTATTTGGCAATTCGCGCTCTCTGGGTTAG 832
QY 805 TTGGATTAACTTCTCCACATCGTCTCTGCTCGGGTATTATGATATCATGTGCTTTTTC 864
Db 833 TTGGAGTTATTTTATTTCTTCTCTTCTCTGGAATTTAGGCGCTTTTTCCTTC 892
QY 865 CAATTACAGAGGTTTTAGCTGTATATTTTACCATGAAAAGTTTTCAAGCTGAGAGGGAC 924
Db 893 CTGTTTACGAAAGTTTGGCTGTAAT-TTTTTTCAAGAAAAATTTCAAGCCGAAAAAGGAG 951
QY 925 TTTCCTTGTCTCTCTCCCTTTGGGGCTT 952
Db 952 TAGCTATTTTCTAGCTCTATGGGGATT 979

Search completed: November 1, 2004, 22:30:29
Job time : 3805.29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 06:20:04 ; Search time 552.64 Seconds
(without alignments)
9900.026 Million cell updates/sec

Title: US-09-913-767-10

Perfect score: 1067

Sequence: 1 agacaagaatgtaggctt.....gaccataagtgagtgttaa 1067

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 253800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	317.8	29.8	1233	16	US-10-424-599-30274
2	284.8	26.7	1146	16	US-10-425-114-10516
3	248.6	23.3	994	16	US-10-424-599-73501
4	194.4	18.2	1733	17	US-10-437-963-57539
5	182.2	17.1	1513	17	US-10-437-963-82061
6	161.8	15.2	2214	16	US-10-424-599-102519
7	145.6	13.6	3387	9	US-09-938-842A-2212
8	145.6	13.6	3387	11	US-09-938-842A-2212
9	136.4	12.8	1373	16	US-10-424-599-31364
10	133.6	12.5	1777	17	US-10-437-963-91425
11	133.6	12.5	1417	16	US-10-424-599-97043
12	132.8	12.4	789	17	US-10-437-963-72452

13	123	11.5	1296	16	US-10-425-114-13687	Sequence 13687, A
14	120.4	11.3	1333	16	US-10-425-114-10574	Sequence 10574, A
15	119.4	11.2	737	16	US-10-425-114-24295	Sequence 24295, A
16	114	10.7	83698	17	US-10-416-898-9	Sequence 9, Appli
17	112.2	10.5	716	16	US-10-424-599-61310	Sequence 61310, A
18	111	10.4	2928	17	US-10-437-963-33800	Sequence 33800, A
19	110.3	10.3	1355	16	US-10-424-599-31363	Sequence 31363, A
20	104.6	9.8	1294	16	US-10-424-599-2524	Sequence 2524, Ap
21	103.8	9.7	997	16	US-10-424-599-92161	Sequence 75161, A
22	98	9.2	792	16	US-10-424-599-9213	Sequence 9213, Ap
23	92.8	8.7	2381	17	US-10-437-963-43860	Sequence 43860, A
24	84.2	7.9	572	16	US-10-424-599-118451	Sequence 118451, A
25	77.2	7.2	1306	17	US-10-437-963-18858	Sequence 18858, A
26	74	6.9	747	17	US-10-767-701-6806	Sequence 6806, Ap
27	72	6.7	592	9	US-09-770-152-44	Sequence 44, Appl
28	67.2	6.3	522	16	US-10-424-599-114931	Sequence 114931, A
29	56.2	5.3	1119	17	US-10-437-963-34510	Sequence 34510, A
30	55.2	5.2	1386	16	US-10-425-114-12485	Sequence 12485, A
31	54.2	5.1	1122	17	US-10-437-963-42424	Sequence 42424, A
32	52.4	4.9	728	17	US-10-767-701-9000	Sequence 9000, Ap
33	52	4.9	1779	16	US-10-425-114-28473	Sequence 28473, A
34	51.6	4.8	2605	17	US-10-437-963-83900	Sequence 83900, A
35	51.4	4.8	486	16	US-10-424-599-74177	Sequence 74177, A
36	47	4.4	774	9	US-09-770-445-895	Sequence 895, App
37	46	4.3	327	16	US-10-424-599-121152	Sequence 121152, A
38	42.2	4.0	41104	9	US-09-816-685-3	Sequence 3, Appli
39	42.2	4.0	41104	16	US-10-639-708-3	Sequence 3, Appli
40	41	3.8	1432	17	US-10-437-963-1897	Sequence 1897, Ap
41	40.6	3.8	6317	15	US-10-204-708-11	Sequence 11, Appl
42	40.6	3.8	6317	15	US-10-311-455-381	Sequence 381, App
43	40.6	3.8	94512	13	US-10-087-192-976	Sequence 976, App
44	39.2	3.7	478	10	US-09-918-995-22681	Sequence 22681, A
45	39	3.7	470	13	US-10-027-632-248589	Sequence 248589, A

ALIGNMENTS

RESULT 1
US-10-424-599-30274
; Sequence 30274, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 30274
; LENGTH: 1233
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12733C.1
; US-10-424-599-30274

Query Match	29.8%	Score 317.8;	DB 16;	Length 1233;
Best Local Similarity	61.3%	Pred No. 1.8e-83;		
Matches 568;	Conservative 0;	Mismatches 347;	Indels 12;	Gaps 3;
Qy	79	TGCGTCTCTACTTCAACAAATCGCGTAAAGAAATTTGTTCTCTACGTTTCTTGAACCTG	138	
Db	2	TGCGTCTCTACTTCAATCCACGGTGCACAAACGAATCTGGCTCTCTAGCTTCTTGAACCTG	61	
Qy	139	CAGGCTTCTCTGTTATCTTCAATCTCTGCTCTTCTTACATACCGCGGCAGCAAGCA	198	
Db	62	CAGCCAATTCGGTCGAGTCTCTCCCAATCAATCAATTTTACATT-----CATAAACG	113	

Qy	199	ACATGTGGGTGATAGTACAAGTTTCTTTCTTATCAAAACCGCGTCTTCTTATCCGCGCTG	258
Db	114	AGTCGTGATCGCTCTTCGGCCCTAATCTC-TATCAAGCCCTCTCTCTCCGCGCTCG	172
Qy	259	TTATTGTAGGCATTCTCTCAGGGTTTGATACTACTTGTATGCATATGGTATAGCTTATC	318
Db	173	CCCTCATCGGACTCTCTACCGGCCTGAGGACTTACTCTACGCTGGCGGTGCTCGCC	232
Qy	319	TTCCAGTTTCTACAGCTGCTCTTATCATTTGCTTACAGTTTAGCTTTTATAGTATCTTCT	378
Db	233	TTCCGGTCTCCACTTTCTCTCTAATCCAAAGCCTCCACCTCGCCTCACGCGCGTCTTCG	292
Qy	379	CATTCTTCAATGGTTAAACATAAGTTCACTCCTTTTACCATCAATGCTGTGTGTTGTA	438
Db	293	CCTTCTCTCTGTCGCGCACAGGTTCAAGCCCTACTCCGTCAACTCGCTCGTCTTCTCA	352
Qy	439	CTGTTGGTCTCGGCTTTTGGGAATGTCATACCGAAACTGATAGCCAGCTTTCATGAGACTC	498
Db	353	CCGTGCGCGCTGTGGTTCTGGCTCTGGCTCCAGCGGGGACCGCCCGCGGTGAGTCGA	412
Qy	499	ACAGCAGTACATAACTGGTTTCTTGATTACTGTAGCAGCAGCTGTTATGTATGCTTTCA	558
Db	413	GTCCGAGTACGTGATTGGTTTGTATTGATTACTTGGCGCTCGGGCGCTGTATGGGTTGG	472
Qy	559	TCTTGCCATTAGTGGAACTGCTTACCAGAAGCTAAGCAAAACCATGAGCTATACCCCTTG	618
Db	473	TTTTTGGCTTGATGGAGTTGGTGTACAAAAGACGGCAGCGTATCACGTACTCTCTGG	532
Qy	619	TGCTCGAGTCCAGTTGATTTTGTGTCCTCTTGGTCTTATTGTGACGCTCATCGGTATGT	678
Db	533	TCATGGAGATTACGCTTGTCTGTGCTTCTTTGCTACCTTATCTGCACCGTTGGAATGA	592
Qy	679	TGATCGCTGGTGATTTCAAGCAGGCCCTTACAAAGAAGCAAGAGAGTTCAAGCTTGGAG	738
Db	593	TAATCAATAATGACTTCAAG--GTGATTCCCGCAGAAGCAAGAGATTTTAAAGCTTGGG	649
Qy	739	AGGCATTGTTCTATGTGTGGCTGTGTTTTTCAGCCATCATATGCAAGGCTCTCTCTCG	798
Db	650	AACAAGTACTACGTTGTGTGTGTGGAGTGCAATAATGTGCAGTTTTTCTCTTGG	709
Qy	799	GAGCCATTGGATTAATCTTCTCCACATCGTCTCTCGTCTCGGGTATTATGATATCAGTC	858
Db	710	GAGCAATAGGGTTATCTTTTGTGCTCTGTTTGTGCTCTTTTGTCCGGTATTATATTTGCTGT	769
Qy	859	TTTTGGCAATTACAGAGTTTTAGCTGTATTATTCTACCATGAAAGTTTCAAGCTTGAGA	918
Db	770	TTCTTCCAGTGCAGGAAGTTTTGGCTGTATTGTATACAAAGAGAGCTTTTCATGCGAGA	829
Qy	919	AGGGACTTTCCTCTGCTCTCTCCCTTTTGGGCTTTGTCTTCTACTTTTATGGTGAGATAA	978
Db	830	AAGGGTTCCTTTGGTGTCTCTCTTTTGGGGGTTTGTCTCTATTCTATGGAGATTA	889
Qy	979	AGTCTGGCGAGGATAAAAGGAAATTC	1005
Db	890	AACAAGACAGGGAAGAAACAGAAATC	916

RESULT 2

US-10-425-114-10516
; Sequence 10516, Application US/10425114
; Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

```

: NUMBER OF SEQ ID NOS: 73128
:
: SEQ ID NO 10516
: LENGTH: 1146
:
: TYPE: DNA
:
: ORGANISM: Glycine max
:
: FEATURE:
:
: OTHER INFORMATION: Clone ID: 700941791_FLI
US-10-425-114-10516

```

Query Match 26.7%; Score 284.8; DB 16; Length 1146;

Best Local Similarity 61.3%; Pred. No. 1.1e-73;

Matches 476; Conservative 0; Mismatches 297; Indels 3; Gaps 1;

Qy	230	TATCAAAACCGCGTCTCTTATACGCGCTGTATTAGGCAATTCCTCTCAGGGTTGTATAA	289
Db	65	TATCAAGACCCCTCTCTTCTCGCCCTCGGCCCTCATCGGACTCCTCACGGCCTCGACGA	124
Qy	290	CTACTTGTATGCATATAGGTATAGCTTATCTTCACAGTTTCTACAGCTGCTCTTATCATATGC	349
Db	125	CTAAGCTTACGCTCGGGCTGCTCGCTTCGGGTCTCCACTTTCCTCTCTAATCCANGC	184
Qy	350	TTACAGATTAGCTTTTATAGCTATCTTCTCATTTCTTCAATGGTTAACATAAGTTCACTCC	409
Db	185	CTCCACCTCGCTTTCACGCGCGTCTTGGCTTCTCTCGTCCGCGCACAGGTTACGCC	244
Qy	410	TTTTACCATCAATGCTTGTGTGTGTGTGACTGTGTGTGCTGCGGTTTGGGAATGCATAC	469
Db	245	CTACTCCGTCAACTCCGTCGTGTCTTCTCACCGTCGCGCTGTGCTGTGCTGTGGCTTC	304
Qy	470	CGAAACTGATAAGCCAGTTTCATCAGACTCACAGCAGTACATAAATCGGTTTCTTGATTAC	529
Db	305	CACGGGGACCGCCCGCGGTGAGTCGAGTCGCCAGTACGTGATGGTTTGTATTGAT	364
Qy	530	TGTAGCAGCAGCTGTTATGTATGCTTTTCATCTTTGCCATTAGTGGAACTTCCTTACCAGAA	589
Db	365	ACTTGGCGTCTCGCGCTGTATCGGTTGCTTTTGGCGGTGATCGAGTGTGTTGTACAAAA	424
Qy	590	AGCTAAGCAAAACATGAGCTATACCTTGTGTGCTCGAGTTCAGGTTGATTTGTGTCTCCT	649
Db	425	GAGCAGCAGCGTATCACGTPACTCTCTGGTTCATCGAGATTACGCTTGTCTTGTGCTTCTT	484
Qy	650	TGCTTCTATTGTTCAGCGTCAATCGGTATGTTTCATCGCTGGTGATTTCGAAGCAGCGCTTACC	709
Db	485	TGCTACTTATTCTGCACCGTTGGAATGATAAATCAATATGACTTCAAG--GTGATTCC	541
Qy	710	AAAAGAACGAAGAGAGTTCAAAGCTTGGAGAGGCAATTGTTCTATGTGGTGGCTGTGTTTC	769
Db	542	CGCAGAAGCAAGAGATTTTAAAGCTCGGGGAAACAAAGTACTACGTTGTGTGGTGGAG	601
Qy	770	AGCCATCATATGGCAAGGCTCTTCTTTGGAGGCACTTGGATTAATCTTCTCCACATCGTC	829
Db	602	TGCAATAAATGTGCAGTTTTTCTTCTTGGGAGCAATAGGGGTTATCTTTTGTGCTCGTC	661
Qy	830	TCTCGTCTCGGTATTATGATATACGTGCTTTTGGCAATTACAGAGGTTTTAGCTGTAT	889
Db	662	TTTGTGTCCGGTATTATAAATGCTGCTTTTCTTCAGTACCGAAGTTTTTGGCTGTTAT	721
Qy	890	ATTCTACCATGAAAAGTTTCAAGCTGAGAAGGCACTTCTCTTGTCTCTCTCCCTTTGGGG	949
Db	722	TGTTATCAAGAGAGCTTTCATCAGAGAAAGGGTTGCTTTGGTGTCTCTCTCTTTGGGG	781
Qy	950	CTTTGTCTCTTACTTTTATGGTGAGATAAAGTCTGGCGAGGATAAAGGAGAATTC	1005
Db	782	GTTTGTGTCTTACTTTATGGAGAGATAAACAAGACAGGAGAAAGAACAGAAATC	837

RESULT 3

US-10-424-599-73501

US 10 424 599 : Sequence 73501, Application US/10424599

; Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 73501
LENGTH: 994
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_37387C.1
US-10-424-599-73501

Query Match 23.3%; Score 248.6; DB 16; Length 994;
Best Local Similarity 57.4%; Pred. No. 6.3e-63;
Matches 510; Conservative 0; Mismatches 369; Indels 10; Gaps 3;

QY 1 AGACAAGAATGGTGAAGGCTCTTGTCATCAATAAATGCATAATCTTAGCCATAGGAAACT 60
Db |||||
QY 110 ACAGAACCATGAAGCCCTCTCTCAGCAAACTGTCTTACTCACCATCGGCACT 169
Db |||||
QY 61 GTGAGGCTCTTTGATATATGCGTCTTACTTCAACAATGGCGGTAAAGGATTTGGTTCT 120
Db |||||
QY 170 CCGGTGGGCCCTCGTCATGCGTCTTACTTCTCCACGGGGCCACCGGCTCGGCTCT 229
Db |||||
QY 121 CTAGGTTCTTGAACCTGAGGCTTTCCTGTATCTTCAATCTCTCTGCTCTTCTTACA 180
Db |||||
QY 230 CCAGCTTCTCGAAACCGGTGGCTTCCCTCTCATGCTCTCTCCCTCGGCTCTCTAC 289
Db |||||
QY 181 TTACCCGGCGCAAGCAACAATGTGG-----GTGATAGTACAAGTTCTTCTTTATCA 234
Db |||||
QY 290 TCGCCGAGGTGCGACCGGCTCGCGCCGCGGAACCGCAAAATTAATCTCAATGA 349
Db |||||
QY 235 AACCCGGCTCTTTATTCGCGGCTGTATTTAGGCAATCTCTCAGGGTTTGATACTACT 294
Db |||||
QY 350 AGCCTCTCTCTCGCGGCTCCACCTTCATCGGAATTTCTACCGGCTCGAGCACTACC 409
Db |||||
QY 295 TGTATGCATATGATAGCTTATCTTCCAGTTTCTACAGTGTCTTATCATTTGCTTAC 354
Db |||||
QY 410 TCTACGCCCTACGCGGTGGGAGGCTTCGGGTCTCCACTTCGCGCTCATCGCAACGC 469
Db |||||
QY 355 AGTTAGCTTTATAGCTATCTCTCATTTCTTATGTTTAAACATAAAGTTTCACTCTTTTA 414
Db |||||
QY 470 AACTCGGTTTACCGGTTCTTCGCGTCTCTCTCGTAGGAGCAAGTTTCAAGCGTACT 529
Db |||||
QY 415 CCATCAATCTGTGTGTGTGTGACTGTGTGTGCTGCGGTTTTTGGGAATGCATACCGAAA 474
Db |||||
QY 530 CCGTAAACCGGTCGTTTTGCTCACTGTGCGGCGCGGCTTTTGGCGCTTCAACCCAGCG 589
Db |||||
QY 475 CTGATAAGCCAGTTTATGAGACTCAAGCAGTACATAACTGGTTTCTTGATTTACGTAG 534
Db |||||
QY 590 GAGACCGTCCCCCGCGAGTCGGTTAAGGAATATGTTATGGCTTTTGTGATGACAGTGA 649
Db |||||
QY 535 CAGCAGCTGTATGATGCTTTCATCTTGCCATTAGTGAACCTTGCTTACCAGAAAGCTA 594
Db |||||
QY 650 TCGCTCGGCAATGTATGATTCATTTTACCTTTGGTGGAGTTGGTGTACAAAATATCA 709
Db |||||
QY 595 AGCAAAACCATGAGCTATACCTTTGCTCGAGTTCAGTTGATTTTGTGCTCTCTTGTCTT 654
Db |||||
QY 710 AACAGCCTCTTACTTCTTTGTCATGGAGATTTCAGTTTCGTTATGTCTTCTCGCCA 769
Db |||||
QY 655 CTATTGTGAGCTATCGGTATGTTTATCGCTGGTGAATTTCAAGCAGGCTTACCAAAAG 714
Db |||||
QY 770 CTCTCTTTTGGCTCTTGGAAATGATCAACAATGACTTTA---AGGTGATTTCGAGGG 826
Db |||||
QY 715 AAGCAAGAGAGTTTCAAGCTTGGAGGCAATGTTCTTATGTGGTGGCTGTGTTTTCAGCCA 774
Db |||||
QY 827 AAGCCAAAATTTGAGCAGCGGAAGGAGTTTACTATGCTGTGTTTGGGAGTGCA 886
Db |||||
QY 775 TCATATGCGAAGGCTTCTTCTTTGGGAGCCATTGGATTAATCTTCTCCACATCGTCTCTCG 834
Db |||||

Db 887 TATTATGGCAGGCTTTTCTTGGGGCGATTGGGCTTATATTTGGGCTGGCTTCTGA 946
QY 835 TCTCGGATTTATGATATACAGTGTCTTTGCCAATTACAGAGCTTTTACG 883
Db 947 TCTAGGAT-TTTTGATTGGGTTGCTACCCGTAACCGAAGTGTGGC 994

RESULT 4
US-10-437-57539
Sequence 57539, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 57539
LENGTH: 1173
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_59340C.1
US-10-437-57539

Query Match 18.2%; Score 194.4; DB 17; Length 1173;
Best Local Similarity 52.3%; Pred. No. 9.5e-47;
Matches 528; Conservative 0; Mismatches 446; Indels 36; Gaps 3;

QY 14 GNAGGCTCTTGATCATAAACTGCATAAATTTAGCCATAGCAAACTGTGGAGGTCCTTT 73
Db 108 GAGCCCGCTCCCTCGTCGCAACTTCGTCCTCATGTGTGTCGCTCGGCGTGGCGCGCT 167
QY 74 GATTATGCTCTCTACTTCAACAATGGCGGTAAAGGATTTGGTTCTCTACGTTTCTTGA 133
Db 168 CTTCTTCGCGCTACTTCTCTCGGCGGGCAACCGAAGTGGCTCTCCAGCTTCTCCA 227
QY 134 AACTCAGGCTTCTCTGTTATCTTCAATCTCTCTCTCTTCTTCAATTAACCGGCGCAG 193
Db 228 GACCGCGGCTGGCGGCTGCTCGCGCGCTCTGCTTCTGTTACTCTCTCACGCGCGG 287
QY 194 AAGCAACAATGGGTGATAGT-----ACAAAGTTTCTTTCT 229
Db 288 CCGGCGCAGGTTGAGGACGACGCGCTGGCGCTGGCGCGGCGCCACGCGCTGTCTCT 347
QY 230 TATCAAAACCGGCTCTTCTTATCGCGCTGTATTCTAGGCAATCTCTCAGGCTTTGATAA 289
Db 348 CATGAGCCACGCTCTCTGTTGGCTCGCGCTCTCGGGCTCATGACCGGCGTGCACA 407
QY 290 CTACTTGTATGCATATGATATAGTTATCTTCCAGTTTCTACAGTGTCTTTATCATTCG 349
Db 408 CTTCTCTACGCTACGCGCTGGCGCTACCTCCCGGTGTCCACTCTCTCATCTCTC 467
QY 350 TTCACAGTTAGCTTTTATAGCTATCTTCTCATCTTCTATGTTTACATTAAGTTTCACTCC 409
Db 468 CACGAGCTGGGCTTTCACGCGGCGCTTCGCGCTGCTGCTGCGCCACGCGGTTCACGCG 527
QY 410 TTTTACCATCAATGCTGTTGTGTTCTGACTGTTCTGCTGCGGTTTTGGGAATCATAC 469
Db 528 GTTCTCGGTGAACCGCGTCTGCTCTCAGGCTCGCGCGGCGCATGCTGGGATGAACGC 587
QY 470 CGAAACTGATAAGCCAGTTTCATGAGCTCACAAAGCAGTACATAACTGGTTTCTTGATTAC 529
Db |||||

FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_63592C.1
US-10-424-599-102519

Query Match 15.2%; Score 161.8; DB 16; Length 2214;
Best Local Similarity 51.3%; Pred. No. 7.1e-37;
Matches 487; Conservative 0; Mismatches 442; Indels 20; Gaps 4;
QY 20 TCTGTGATCATAACTGCATAATCTAGCCATAGAAACTGTGGAGGTCCCTTGTGATTAT 79
DB 416 TCTGTGACACTCAGCATAGCCTTCTTATAGTTGGCCAATCTGCTGCTGTATCCTGG 475
QY 80 GCGTCTCTACTCAACAATGGCGGTAAAGATTGGTCTCTACGTTTCTTGAAGACTGC 139
DB 476 AAGATTTTATGATCAGGGTGAATAATAGTAATGGATGGCTACTCTAGTTCAAACTGC 535
QY 140 AGCGTTTCTGTATCTTCAATCTCTGCTCTCTTCTTACATTAACCGGCGCAGAGCAA 199
DB 536 TGCCTTCCCGAATCTTGTTCATTCATATTT-TACAATTCCTTCACCTCCAGAGGCTTCAA 594
QY 200 CAATGTGGGTGATGATGACAAAGTTTCTTCTTATCAAAACCGCGTCTTCTTATCGCGCTGT 259
DB 595 C-----TTCTGCTTCACTCCATCAAAATTAATCTTTCATATATTTGG 640
QY 260 TATGTAGGATTTCTCTCAGGGTTTGATTAATACTTCTGTATGCAATGATAGTATATCT 319
DB 641 TCT--TTGGAGTCTTAATTTGCTGCTGACAATATGATGTACTCCACTGGACTCTTATACCT 698
QY 320 TCCAGTTTCTACAGTCTCTTATCAATCTGCTTACAGTTAGCTTTTATAGCTATCTTCTC 379
DB 699 CTCGGCTTCTACCTATTCGCTGATTTGTGCATCACAGTTAGCTTTTAATGCAAGTTTCTC 758
QY 380 ATCTTTCATGGTTAAACATAAAGTTTCACTCTTTTACCATCAATGCTGTGTGTGTGAC 439
DB 759 ATATTTTATCAATCTCAAAAGTTTCACTGCTTGTATTAATACTCTACAGTGGTCTCAC 818
QY 440 TGTGTGTGTGCGGTTTGGGAATGATACCGAAACTGATAAGCCAGTTTCATGAGACTCA 499
DB 819 TTATCTGTGCACTCTTGTCTTAAACGAAGACACAGATGAACATCTGGTTTCTCCAA 878
QY 500 CAAGCAGTACATAAATCTGTTTCTTGATTAATCTAGCAGCAGCTGTTATGATGCTTTCAT 559
DB 879 GGAAGATGATATTTAGTTTCTTATGATACCTTTGGAGCTTCTGCACTGCTCTCTCTTT 938
QY 560 CTTGCCATTAAGTGGAACTTGTCTTACCAGAAAGCTAAGCAAAACCATGAGCTATACCTTGT 619
DB 939 GCTTTCCTCATGCAGCTGACCTTTGAGAGGTTCTGAAGAGGAAACAATTTCTGTGTGT 998
QY 620 GCTCGAGTTCCAGTTGATTTTGTGCTCCTTGTCTTATTTGTGAGCGTCAATCGGTATGTT 679
DB 999 TTTGGAATGCAAACTACACATCAATTCGTTGCTCTGCTGCTCTGTCATAGGCTATT 1058
QY 680 CATCGTGTGTATTCAGCAGGCTTACCAAGAGCAGAGAGTTCAGGTTGGAGA 739
DB 1059 TCGAAGTGGGAATG---CGTACTTTGATGAGAAATGGAGGGTTTTCAGAAAGGATA 1115
QY 740 GGCATTTGTTCTATGTGGTGGCTGTGTTTTCAGCCATCATATGCAAGGCTTCTTCTGGG 799
DB 1116 TGTGTCTTATGTTATGACTTTGGTTTGGACTTCAATAGCTGGCAGGTAATGCTGTGG 1175
QY 800 AGCCATTGGAATTAATCTTCCACATCGTCTCTCGTCTCGGGTATTAATGATATCAAGTGT 859
DB 1176 TGTGTGTGGCTTGATCTTCTAGTGTCTTCTCTACTCCAATGTTTAAAGCACAGTTTC 1235
QY 860 TTTGCCAATTACAGAGGTTTATGCTTATATTTTACCATGAAAGTTTCAAGCTGAGAA 919
DB 1236 TTTAGCCGTAACCTTATGCTGCTGTATAGTTTTCATGATAAGATGAATGGGGTGAA 1295
QY 920 GGGACTTCTCTGTGCTCTCTCCCTTTGGGCTTTTGTCTCTTACTTTTAT 968
DB 1296 GATAATTTCTATGCTTTTGGCTCTATGGGTTTTCCTCTTATATTTAT 1344

RESULT 7
US-09-938-842A-2212
Sequence 2212, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2212
LENGTH: 3387
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2212
Query Match 13.6%; Score 145.6; DB 9; Length 3387;
Best Local Similarity 49.1%; Pred. No. 6.1e-32;
Matches 475; Conservative 0; Mismatches 484; Indels 9; Gaps 3;
QY 1 AGACAAGAAATGGTGAAGGCTCTTGTGATCATAAATGCAATATCTAGCATAGGAACT 60
DB 2354 AGAAGCTGTAGAGGTGGCTCCGTGCTCCATATACGAATCTTGTCTATCTTCTGCCAAC 2413
QY 61 GTGAGGTCTCTTGTATATGCGTCTCTACTTCAACAATGGCGGTAAAAAGGATTGTTCT 120
DB 2414 CACTTGTACAGTTCTGGGTAGACTGTACTATGAAATGGAGGAAAGACATATGTGG 2473
QY 121 CTACGTTCTTCAAACTGCAGGCTTCTGTTATCTTATCTTCTCTCTCTCTCTTACA 180
DB 2474 TAACACTTCTTCAACTCAATGGCTTCCCTGTACTGATCTCTTCGCTTCTTCTCGAA 2533
QY 181 TTACCCGCGCAGAGCAACAATGGGTGATGATACAAAGTTTCTTCTTATCAAAACCGC 240
DB 2534 TCAGCAACCCAAATCAACAGATACAAATTTCACT--CAGTCCCTTCTTCCACCCCT 2591
QY 241 GTCTTTCTTATCGCGCTGTTATTTAGGCAATCTCTCAGGGTTTGATTAATCTTGTATG 300
DB 2592 TGCATCGGTTTACTTTGTGCACT---GGACTGCTAGTGTCTGCTTATGCTTATTTGCTG 2647
QY 301 CATATGTTATAGCTTATCTTCCAGTTTCTACAGTGTCTTATCATCTTCTTCTTACAGTTAG 360
DB 2648 CAGTTGGGTGCTTTTACTTACAGTCTCTACTTTTCTCCCTCATCTTGGCCTCACAGTTGG 2707
QY 361 CTTTTATAGCTATCTTCTCATCTTTCATGGTTTAAACATAAAGTTCACTCTTCTTACATCA 420
DB 2708 CTTTCACTGCCTTTTCTCATATTTCTTAACTCCGAAAGTTTCACTCTTCTTGTATGTA 2767
QY 421 ATGCTGTTGTTGTTGACTGTTGTGCTGCGGTTTGGGAATGATCATACCGAAACTGATA 480
DB 2768 ATTCTTGTGTTTCTTACGGTTTCTCTGCTCCCTCTCTGCTGCTCAACTGATTGAGAAA 2827
QY 481 AGCCAGTTTCATGAGACTCACAAGCAGTACATAAAGTTTCTTGTATTTACTGTAGCAGAG 540
DB 2828 ACACAACAAATGATCTAGAGTACAAATATGATTTGGGTTTCACTGTACCATTTGTGCTT 2887
QY 541 CTGTTATGATGCTTTTCACTTCTTCCATTTAGTGGAACTTGTCTTACAGAAAGCTAAGCAA 600
DB 2888 CCGCTGGGATGGATTGGTACTTCTCTGATACAACTGCTCTTTCAGGAAAGTTTTCACGA 2947
QY 601 CCAAGAGCTATACCCCTTGTGCTCGAGTTCCAGTTGATTTTGTGTTCTCTTCTTATTTG 660

2948	Db	AGCATACATCTCAGCAGTCTCGGAGCTTGGCCAAATTACCACTCTCTAGTTGCCAACTTG	3007
661	Qy	TCAGCGTCATCGGTATGTTTCATCGCTGGTGAATTTCAAGCAGGCGCTTACCAAAAGAAGCAA	720
3008	Db	TGGTACTCATGAGTGTGTTGCAAGTGGAGT---GGAGNACTCTGCCAAGTGAGATGA	3064
721	Qy	GAGAGTTCAAGCTTGGAGAGCAATTGTTCTATGTGGTGGCTGTGTTTTCAGGCCATCATAT	780
3065	Db	GAAACTACAACTGGGGAAGTGTCATATCTTTGACTTTGGCGCTCAGCAGCTATTTTCT	3124
781	Qy	GGCAAGGCTTCTCTCGGAGCCATTGGATTAATCTTCTCCAATCGTCTCTCGTCTCGG	840
3125	Db	GGCAAGTATACACTGTGGTTGTGTGGGATTAATCTTCGAGTCTTCTCTGTGTCTCCA	3184
841	Qy	GTATTATGATATCAGTGCCTTTGGCCAATTACAGAGGTTTTAGCTGTATATCTTCAACATG	900
3185	Db	ATTCCTAACTGCTGTGGGACTGCCTATAGTTCAGTTGTAGCAGTGATAGTTTTCCATG	3244
901	Qy	AAAAGTTTCAAGCTGAGAAGGCACTTCTCTTGTCTCTCCCTTTGGGGCTTTGCTCTTT	960
3245	Db	ATAAGATGGATGCATCAAGATTTTCTCCATCATTTTAGCTATCTGGGGCTTCTTTTCAT	3304
961	Qy	ACTTTTAT	968
3305	Db	TTGTCTAT	3312

RESULT 8

```

US-09-938-842A-2212
; Sequence 2212, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2212
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2212

```

Query Match	13.6%	Score 145.6	DB 11	Length 3387
Best Local Similarity	49.1%	Pred. No. 6.1e-32		
Matches 475; Conservative 0;	Mismatches 484;	Indels 9;	Gaps 3;	
Qy	1	AGACAAGATGGTGGAAGGCTCTTCGTGATCATAAACTGCATAATTTCTAGCCATAGGAAC	60	
Db	2354	AGAACTGTAGAGGTGGCTCGGTGTCCTCATATAGCAATCTTTGTGCATCTTCTGCCAAC	2413	
Qy	61	GTGGAGTCTCTTTGATTATGGCTCTCTACTTCAACAATGGCGGTAAAGATTGTGGTTCT	120	
Db	2414	CACCTGTCTACAGTTCTGGGTAGACTGTACTATGAATAATGGAGGAAAGCACATATGTGG	2473	
Qy	121	CTACGTTTTCTTGAACCTGCAGGCTTTCCTGTTACTTCTCATTTCCCTGCTCTTCTCTTACA	180	
Db	2474	TAAACATCTTTCAACTCACTCATTTGGCTTCCTGTACTGATTCTCTTCGCGTCTCTTTTCGAA	2533	
Qy	181	TTATCCCGCGCAGAGCAACAATGTGGTGATGTACTACAAGTTCTTTCTTATCAAAACCGC	240	
Db	2534	TCAGGCAACCCAAATCAACAGATACAAATTTCACT--CAGTCCCGTTCTTCTCACACCTT	2591	

Qy	241	GTCTTTCTTATCGCGCTGTTTATTGTAGGCAATCTCTCAGGGTTTGATACTACTGTGATG	300
Db	2592	TGCATCGGTTTACTTTGTGCACT---GGACTGCTAGTGTCTGCTATATGTTTGTCTG	2647
Qy	301	CANATAGTATAGCTTTACTTCCAGATTCTTACAGCTGCTTTATCATTTGCTTTCACAGTTAG	360
Db	2648	CAGTTGGTTGCTTTACTTACCAGTCTCTACTTTTCTCCCTCATCTTGGCCCTCACAGTTGG	2707
Qy	361	CTTTTATAGCTATCTTCTCATTTCTTCATGTTTAAACATAAAAGTTTCACTCCTTTTTACCATCA	420
Db	2708	CCTTCACTGCCCTTTTCTCTCATATTTCTTAACTCGCAAAAGTTTCACTCCTTTTGATAGTCA	2767
Qy	421	ATGCTGTTGTTGTTGACGTGTTGGTGCTGCGGTTTTGGGAATGCATACCGAAACTGATA	480
Db	2768	ATTCTTTGTTTCTCTTACGGTTTTCTCTGCCCTCTCTCGTGGTCAACACTGATTCAGAAA	2827
Qy	481	AGCCAGTTTCATGAGACTCACAAACAGTACATAACTGGTTTTCTTGATTACTGTAGCAGCAG	540
Db	2828	ACACAACAAATGATCTTAGAGTACAATATGTGATTTGGGTTTCATCTGTACCATTTGGTGCTT	2887
Qy	541	CTGTTATGTATGCTTTTCATCTTGCCATTAGTGGAACTTGCTTACCGAAAGCTTAAGCAAA	600
Db	2888	CCGCTGGGATTGGATTGGTACTATCTCTGATACAACTGCTCTTCAGGAAAGTTTTCACGA	2947
Qy	601	CCATGAGCTATACGCTTGCTCGAGTTCCAGTTGATTTTGTGTCTCCTTGTCTTCTATTG	660
Db	2948	AGCATACATCCTCAGCAGTCTCGAGCTTGCCCAATTACCAAGTCTCTAGTTTGCACATTGTG	3007
Qy	661	TCAGCGTCATCGGTATGTTTCATCGCTGGTGATTTCGAAGCAGGCTTACCAAAAAGAACAA	720
Db	3008	TGGTACTCATPAGACTGTTTGGAAAGTGGAGAGT---GGAGAACTCTGCCAAAGTGAGATGA	3064
Qy	721	GAGAGTTTCAAAGCTTTGGAGAGGCATGTTCTTATGTGGTGCGTGTGTTTTTCAGCCATCATAT	780
Db	3065	GAAACTACAAACTGGGAAAGTGTATATCTTTGACTTTGGCCTCAGCAGCTATTTTCT	3124
Qy	781	GGCAAGCTTCTTCTTGGGAGCCATTGGGATTAATCTTCTCCACATCGTCTCTCGTCTCGG	840
Db	3125	GGCAAGTATACACTGTTGGTTGTGTGGGATTAATCTTCGAGTCTTCTCTGTGTGTTTCCCA	3184
Qy	841	GTATTATGATATCAGTGTCTTTTCGCCAATTACAGAGTTTTTGTGTTTATATTCTACCATG	900
Db	3185	ATTCCATACTGCTTGGGACTGCCCTATAGTTCCAGTTGTAGCAGTGAATAGTTTTCATG	3244
Qy	901	AAAAGTTTCAAGCTGAGAAGGACTTTCTCTGCTCTCTCCCTTTGGGGCTTTGTCTCTT	960
Db	3245	ATAAGATGGATGCATCAAAAGATTTTCTCCATCATTTTAGCTATCTGGGGCTTCTCTTCAT	3304
Qy	961	ACTTTTAT	968
Db	3305	TTGTCTAT	3312

RESULT 9

US-10-424-599-31364
 ; Sequence 31364, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 31364
 ; LENGTH: 1373
 ; TYPE: DNA
 ; ORGANISM: Glycine max


```
Db 642 AACCGCTGCAATGACCAATCAATGCCCTGGACTCGAGTTCAGACAGATATCCAGCAATTAG 701
Qy 497 TCACAAGCAGTACATAAATCGGTTTCTTGATTACTGTAGCAGCAGCTGTTATGATGCTTT 556
Db 702 TGACAGTGAATACATCATGGGATTTGTATGGGATGTTTAGGATCTGCTTTCCATGGGCT 761
Qy 557 CATCTTGCCATTAGTGGAACTTGTTCACAGAAAGCTAAGCAACCATGAGCTATACCCT 616
Db 762 TATTTTCGCTCTCTCAGAGCTCGTCTTGTGAAGTGTCTCGAAGAAGATCCTTTATCGT 821
Qy 617 TGTGCTCCAGTTCAGATTGATTTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 676
Db 822 TGTTCGAGGACGAAGTCAATGTTCTCTATTTGCAATTTCTGTTTACCACCTGTAGGGAT 881
Qy 677 GTTCATCGCTGCTGATTTCAAGCAGGCTTACCAAAAGCAAGAGAGTTCGAAGCTTG 736
Db 882 GATTGTGAGTGTGATTT---TCAAGGATGACATGAGGCTACCATTTTGAAGGTGG 938
Qy 737 AGAGCAATGTTCTATGTGGTGTGTGTTTTCAGCCATCATATGGAAGGCTTCTTCTT 796
Db 939 TAGAAGTGTCTTATATCTGTTATCATTTTGGGTGCAATCACTTTTCAGCTGGGGTTCT 998
Qy 797 GGGAGCCATTGGAATTAATCTTCCACATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 856
Db 999 GGGGGGCACTGTATAAATTTTCTTGGGCTCTACTGTCTAGCAGGTGTGCTTAATGCAGT 1058
Qy 857 GCTTTTGGCAATTACAGAGGTTTATAGCTGTTATATTTCTACCATGAAAGTTTCAAGCTGA 916
Db 1059 AAGAACCCATTAACAGCAATTCAGCTGTTATATCTGCTTAAGGACCTATGAGTGGTTT 1118
Qy 917 GAAGGACTTCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 970
Db 1119 CAAGATCCTCTCCCTAGTGAATCACCTTTTGGGGAATTTGGCTCATATATTTAAG 1172
```

RESULT 15

```
US-10-425-114-24295
; Sequence 24295, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 24295
; LENGTH: 737
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3608-022-E3_FLI
US-10-425-114-24295
```

```
Query Match 11.2%; Score 119.4; DB 16; Length 737;
Best Local Similarity 52.0%; Pred. No. 1.5e-24;
Matches 298; Conservative 0; Mismatches 266; Indels 9; Gaps 1;

Qy 20 TCTTGTGATCATAAATCTAGCCATAGGAACTGTGGAGTCTCTTTGATTAT 79
Db 165 TCTCGTAGCGCTCAACTCGGGATGCTCGCGGTGGAGCCATCGCGGGCGGCTCCTGAG 224
Qy 80 GCGTCTCTACTTCAACAATGGCGGTAAAGGATTTGGTTTCTCTAGCTTTCTTGAACTGC 139
Db 225 CCGGCTCTACTTACAGAGGGGGGACCCGCGAGTGGCTGTCCGGTGGCTGGAGACGG 284
Qy 140 AGGCTTTCCTGTTATCTTCAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 199
```

```
Db 285 CGCGTGGCCGCTGCTGCTGATTCCCGTGGCGCGTCTGTTGCGCCGCGCGCGGA 344
Qy 200 CAATGTGGGTGATAGTACAAGTTTCTTTCTTATCAAAACCGGCTCTTCTTATCGCGGCTGT 259
Db 345 CAGGG-----GCGCCCGGCTCTGCTCGCGCCCGCGGATCCTGCTGGCGCGCGC 395
Qy 260 TATTGTAGGCATTCTCTCAGGGTTTGATACTACTTGTATGCATATGGTATAGCTTATCT 319
Db 396 GGGCTCGGCTGGCCACGGGCGGAGACTTCACTACGCCCTTCGGGCTCGGCTACCT 455
Qy 320 TCCAGTTTCTACAGCTGCTCTTATCATTTGTCTTCAAGTTAGCTTTTATAGCTATCTTCTC 379
Db 456 CCGGTTGTCCACGTCCGCGATCCTGATCTCGACGAGCTGGCGTTACGGTGTCTTTCGC 515
Qy 380 ATTCTTTCATGGTTAAACATAAAGTTTCACTCTTTTACCATCAATGCTGTGTGTTGTTGAC 439
Db 516 GTTCTCTGTTGGTGGCGAGCGCTGACGGCGCGTCCGTGAACCCGTTGGCGCTGCTGAC 575
Qy 440 TGTGGTGTCTCGGTTTGGGAATGCATACCGAACTGATAAGCCAGTTTCATGAGACTCA 499
Db 576 GGTGGGCGCGTGGTGTGGGCTGCGAGCTCTCGTCGACCGGCCCGCGGCGTCAACAG 635
Qy 500 CAAGCAGTACATAAATCGTTTCTTGTATGTTAGTACGACAGCTGTGTATGATGCTTTTCA 559
Db 636 GGGCCAGTACTGGCTGGGCTTCGTGTGACGCTCGCGCGCGCGCTCTACGGCCTCGT 695
Qy 560 CTTGCCATTAGTGAACCTTGTTCACAGAAAGC 592
Db 696 CTTGCCCTCGTGGAGCTCGCTACAGGCGGC 728
```

Search completed: November 1, 2004, 23:40:13

Job time : 555.64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:48:03 ; Search time 98.4346 Seconds
(without alignments)
7704.726 Million cell updates/sec

Title: US-09-913-767-10
Perfect score: 1067
Sequence: 1 agacaagaatggtgaaggct.....gaccataagtgaagttaa 1067

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A-COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B-COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A-COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B-COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS-COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59.4	5.6	7218	1	US-08-232-463-14
2	43	4.0	1141	4	US-09-806-708B-22
3	41.6	3.9	289	3	US-09-007-005-17
4	41.6	3.9	289	3	US-09-244-796-17
5	40.6	3.8	6317	4	US-10-204-708-11
6	37.4	3.5	933	4	US-09-248-796A-13724
7	37.2	3.5	1055	4	US-09-806-708B-23
8	37.2	3.5	1664976	4	US-08-916-421B-1
9	37.2	3.5	1664976	4	US-09-692-570-1
10	37	3.5	1308	4	US-09-270-767-12733
11	37	3.5	4182	1	US-07-973-257-1
12	36.2	3.4	1738	3	US-09-202-712-18
13	35.8	3.4	399	4	US-09-621-976-8976
14	35.8	3.4	832	4	US-09-621-976-2813
15	35.4	3.3	534	4	US-09-248-796A-751
16	35.4	3.3	832	4	US-09-621-976-2813
17	35.2	3.3	1485	3	US-09-134-001C-1638
18	35	3.3	580073	4	US-08-545-528D-1
19	34.6	3.2	618	4	US-09-248-796A-1121
20	34.6	3.2	1100	3	US-08-642-807A-32
21	34.4	3.2	2007	4	US-09-614-221A-370
22	34.2	3.2	717	3	US-08-998-416-602
23	33.6	3.1	703	3	US-08-998-416-178
24	33.6	3.1	723	3	US-08-998-416-952
25	33.6	3.1	2139	4	US-09-248-796A-5542
26	33.6	3.1	2733	4	US-09-107-532A-2699
27	33.4	3.1	534	4	US-09-248-796A-1197

28	33.4	3.1	1068	4	US-09-328-352-2430
29	33.4	3.1	1167	3	US-09-308-003-6
30	33.4	3.1	1539	4	US-08-956-171E-562
31	33.4	3.1	1539	4	US-08-781-986A-562
32	33.4	3.1	5163	3	US-08-700-651-1
33	33.4	3.1	5163	3	US-08-928-361B-4
34	33.4	3.1	5163	3	US-09-588-995A-4
35	33.4	3.1	5318	3	US-08-700-651-2
36	33.4	3.1	5318	3	US-08-928-361B-3
37	33.4	3.1	5318	4	US-09-588-995A-3
38	33.2	3.1	1233	3	US-08-858-207A-120
39	33.2	3.1	1863	1	US-09-198-284-4
40	33.2	3.1	1863	2	US-08-987-122-4
41	33.0	3.1	3010	4	US-08-961-527-25
42	33.2	3.1	3534	3	US-09-134-001C-2269
43	33.2	3.1	5191	4	US-09-809-665A-41
44	33	3.1	3259	3	US-09-318-448-23
45	33	3.1	8607	4	US-10-204-708-71

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Hardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMM
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZapt-F1s
US-08-232-463-14

Query Match 5.6%; Score 59.4; DB 1; Length 7218;

[illegible]


```
;/ NAME/KEY: misc feature
;/ LOCATION: (231980)..(231980)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (234187)..(234187)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (234220)..(234220)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (234814)..(234814)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (309398)..(309398)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (309418)..(309418)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (312837)..(312837)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (312993)..(312993)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (319226)..(319226)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (559167)..(559167)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (559241)..(559241)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (600992)..(600992)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (622708)..(622708)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (657081)..(657081)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (657203)..(657203)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (674435)..(674435)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (682442)..(682442)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (713652)..(713652)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (741684)..(741684)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (779455)..(779455)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (779676)..(779676)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (855539)..(855539)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (871619)..(871619)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1084830)..(1084830)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
```

```
;/ LOCATION: (1096846)..(1096846)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1119881)..(1119881)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1130881)..(1130881)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1310988)..(1310988)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1313224)..(1313224)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1349473)..(1349473)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1349491)..(1349491)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1470091)..(1470091)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1569020)..(1569020)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1602912)..(1602912)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1603734)..(1603734)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1637998)..(1637998)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1664854)..(1664854)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ US-08-916-421B-1
```

Query Match 3.5%; Score 37.2; DB 4; Length 1664976;

Best Local Similarity 43.8%; Pred. No. 19;

Matches 162; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

```
Qy 213 AGTACAAGTTCTTTCTTATCAACCGCTCTTCTTATCGCGCTGTTATTCTAGGCATT 272
Db 1077557 ACTACTGGATTGGTATTTAAAGAGAAGTATTATAAGCATCTGCAGTTATTGGGTTAATT 1077498
Qy 273 CTCTCAGGTTTGATAACTTTGTATGCATATGGTATAGCTTATCTTCCAGTTTCTACA 332
Db 1077497 CTATTGTTGTTATTTTATTCGTGTTTTTACAGCATATAACTCAATTGGAGCAATTTTA 1077438
Qy 333 GCTGCTTTATCATGTTCTCAGTAGCTTTTATAGCTATCTTCTCATCTTCATGGTT 392
Db 1077437 TACGTTGGAATCTTAGTTGGAATATTTTATTATTAGTTAAATTTTATAGATAAAACCTTA 1077378
Qy 393 AACATAAAGTTTCACTCTCTTTTACCATCAATGCTGTTGTTGTTGTTGTTGTTGTTGTTG 452
Db 1077377 AAATTAAATATCATACAGTCAAGATTGATGATTGTTGTTGTTGTTGTTGTTGTTGTTG 1077318
Qy 453 GTTTTGGGAATGCATACCGAAACTGATAAGCCAGTTTCATGAGCTTCACAAGCAGTACATA 512
Db 1077317 GCTTCTGCAACAACATTATTGGAATAGGAGTTTATGTTTATGTTTATGTTTATGTTTAT 1077258
Qy 513 ACTGTTTCTTGATTACTGTAGCAGCAGCTGTTTATGTTGTTTCTCATCTTCTGTCATTAGTG 572
Db 1077257 AGATTTTGTATGGAACCTTTTGGAGTTTATGCTTTTATTCCTTTTAAAGCTCTTATAATT 1077198
Qy 573 GAACCTTGCTT 582
Db 1077197 TTATTAGCTT 1077188
```

```
RESULT 9
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
```


;
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM/PC or Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/973,257
; FILING DATE: 19921109
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: No. 5378820e
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4182 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE:
; ORGANISM: Mycoplasma Gallisepticum
; STRAIN: S6
; US-07-973-257-1

Query Match 3.5%; Score 37; DB 1; Length 4182;
Best Local Similarity 53.9%; Pred. No. 0.9;
Matches 76; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 319 TTCAGTTCTACAGCTGCTTATCAATGCTTACAGTTAGCTTTTATAGCTATCTTCT 378
Db 4163 TTTCAAAACAGCTAATCCACTATCAAGTATAACTTTTAAAGTTTATAGAAATATTC 4104
QY 379 CATTCCTTCATGGTTAAACATAAAGTTCACTCCTTTTACCATCAATGCTGTGTGTTGTA 438
Db 4103 TAICTTAATATATTACCTTAATTAATTCATTTCTTAGTTGGTCTGGAGCTTTTGTT 4044
QY 439 CTGTTGGTCTGCGGTTTGG 459
Db 4043 TAGCTGGTCTGTTGGTTTG 4023

RESULT 12
US-09-202-712-18
; Sequence 18, Application US/09202712
; Patent No. 6265637
; GENERAL INFORMATION:
; APPLICANT: Coupland, George M
; APPLICANT: Schaffer, Robert J
; TITLE OF INVENTION: Genetic control of flowering
; FILE REFERENCE: 620-55
; CURRENT APPLICATION NUMBER: US/09/202,712
; CURRENT FILING DATE: 1999-01-11
; EARLIER APPLICATION NUMBER: PCT/GB97/01676
; EARLIER FILING DATE: 1997-06-23
; EARLIER APPLICATION NUMBER: GB 9613132.1
; EARLIER FILING DATE: 1996-06-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 1738
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-202-712-18

Query Match 3.4%; Score 36.2; DB 3; Length 1738;
Best Local Similarity 53.1%; Pred. No. 0.98;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 212 TAGTACAAGTTCTTTCTTATCAACCGGCTCTTCTTATCGCGCTGTTATTGTAGGCAT 271

Db 1018 TAAACAGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1077
QY 272 TCTCTCAGGGTTTGAATAACTACTTGTATGATATGGTATAGCTTATCTTCCAGTTTCTAC 331
Db 1078 TCTCTCTGTTTTTAAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1137
QY 332 AGCTGCTCTTATCATCTTCTTCACAG 356
Db 1138 ATTTCCGGGAACGATGACTTCTCCG 1162

RESULT 13
US-09-621-976-8976/c
; Sequence 8976, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8976
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match 3.4%; Score 35.8; DB 4; Length 399;
Best Local Similarity 12.5%; Pred. No. 0.59;
Matches 35; Conservative 134; Mismatches 106; Indels 6; Gaps 1;
QY 669 ATCCGTATGTTCACTCGGTGGTGATTTCAAGCAGGCCTTACCAAGAGAGAGAGAGTTC 728
Db 314 AKMKRKTTKMYMKGGKGGSTYAMRSRGSTGRWSYRRAMWRGSKSWGGSYRM 255
QY 729 AAGCTTCGAGAGCAATGTTTCTATGTGTGGTGTGTTTTCAGCC-----ATCATATGG 782
Db 254 AGYRSSRWSYSAMWRKKMTCKGRSSGSRSTGYVWYMKSWCTSRKWMYKKRRK 195
QY 783 CAAGGCTTCTTCTGGAGGCATTTGATTAATCTTCTCCACATCGTCTCTCGTCTCGGT 842
Db 194 KWRKCTSTKRTCYRGSTYKCAAYTKRRKTRWTTTYYKYSMSMKTRMKTYWT 135
QY 843 ATTATGATATCAGTGCTTTTGCCCAATTACAGAGGTTTGTAGCTGTATATTTACCATGAA 902
Db 134 KRWMTRTKWTCTMCWKCTTYWAGTMYRYYRYVYAKRWSKRCSTWSTTCYCKMYM 75
QY 903 AAGTTTCAAGCTGAGAAGGAGCTTTCTCTTCTCTCTCTCCCT 943
Db 74 AKCWSYWSMSMMKMGKSMWKKWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 34

RESULT 14
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match
Best Local Similarity 3.4%; Score 35.8; DB 4; Length 832;
Matches 25; Conservative 103; Mismatches 85; Indels 0; Gaps 0;

QY 816 TTCACATCGTCTCTCGTCTGGGTATATGATATCAGTGTCTTTTGGCAATTACAGAG 875
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
7 KYTTWAKCWTWKWSWSYMYWKYMKTYWRERKKKAWWKYKWTWTWYRYAMWG 66
QY 876 GTTTTAGCTGTTATATTACCATGAAAGCTTCAAGCTGAGAAAGGACCTTCTCTTGCT 935
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
67 TYKKKAMCRTKTKKKKKKGYMMWYWGWRRSYMMWTRTWGYAYRSMYMWRYRCWKK 126
QY 936 CTCCTCCTTTGGGGCTTTGTCTCTTACTTTTATGATGAGATAAGTCTGGCGAGGATAAA 995
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
127 KAYRKTTCYSSKGTWKKRWKKAWTWWKKTYWAAATRYMMWMCWTKWRASWYWCW 186
QY 996 AGGAGAAATTCAGCAGGAGAGAGTCAAGAGACA 1028
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
187 WGRKRWSTWRKRSYASARSARAKRCCYSCSWG 219

RESULT 15
US-09-248-796A-751/c
; Sequence 751, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 751
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-751

Query Match
Best Local Similarity 3.3%; Score 35.4; DB 4; Length 534;
Matches 84; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 219 AGTTTCTTTCTTATCAAAACGCGTCTTCTATCGCGCGTGTATGTTAGGCAATTTCTCTCA 278
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
518 AATTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 459
QY 279 GGGTTTGATAACTACTTGTATGATATGATATGATATGATATGATATGATATGATATGATATG 338
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
458 GCCAATCTTTTCATACCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 399
QY 339 CTTATCATTCATACAGTAGCTTTTATAGCTATCTCTCTCATTC 383
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
398 ATTCGTTGATCTTAGCAGATCTTTTTCATCTCTCAACTCTTTC 354
```

Search completed: November 1, 2004, 22:41:08
Job time : 104.435 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2004, 05:45:29 ; Search time 4253.75 Seconds
(without alignments)
10228.408 Million cell updates/sec

Title: US-09-913-767-5
Perfect score: 1194
Sequence: 1 tcatgagatataataacat.....tgtgatcaagcatatttcc l194

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hc: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gse1: *
9: gb_gse2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1151.6	96.4	1340	CNS0A29J	BX827536 Arabidops
2	1091.2	91.4	1359	CNS0A2R3	BX827477 Arabidops
3	941.6	78.9	1431	CNS0A2KR	BX829081 Arabidops
4	909.2	76.1	1375	CNS0A2S8	BX827187 Arabidops
5	903	75.6	1312	CNS0A2U8	BX827470 Arabidops
C 6	491	41.1	699	B77316	B77316 T3213TF TAM
7	442.6	37.1	709	BH498028	BH498028 BOHOK68TF
8	437.4	36.6	691	BZ087060	BZ087060 lkh10d12
9	396	32.3	558	AV828990	AV828990 AV828990
C 10	334.8	28.0	446	AI996743	AI996743 70168184
11	317	26.5	1519	CNS0A3B5	BX826795 Arabidops
12	312.8	26.2	1323	CNS0A36P	BX826811 Arabidops
C 13	299.6	25.1	498	AU226302	AU226302 AU226302
14	239.8	20.1	922	BZ967767	BZ967767 PUDGUS6TD
C 15	238	19.9	428	AU227210	AU227210 AU227210
C 16	235.8	19.7	424	BP575101	BP575101 BP575101
17	234.8	19.7	739	CO105307	CO105307 GR_Eb003
C 18	231.2	19.4	568	BM176948	BM176948 saJ74h01
C 19	229.8	19.2	782	CO105310	CO105310 GR_Eb003
C 20	228.6	19.1	770	CG344371	CG344371 OG4AG15TC
C 21	228.4	19.1	822	BZ491815	BZ491815 BONFU42TF
C 22	224.6	18.8	837	CB292743	CB292743 UCRC501.0
C 23	224.4	18.8	351	AV441569	AV441569 AV441569
C 24	224.2	18.8	444	AV439952	AV439952 AV439952

25	220	18.4	806	4	BI310188
26	219.6	18.4	666	5	BQ157452
27	219.2	18.4	1374	3	CNS0ABEV
28	213	17.8	401	7	T42237
29	212.6	17.8	627	5	BQ155273
30	212.6	17.8	729	8	BH470210
31	211.4	17.7	905	5	BX927581
32	208.2	17.4	768	1	AJ795195
33	208	17.4	780	6	CD485871
34	200.4	16.8	633	1	AI777352
C 35	199.8	16.7	426	5	BP666513
C 36	198.4	16.6	562	7	CK759270
37	195	16.3	668	5	BQ704880
38	190.6	16.0	630	4	BG525645
39	190.6	16.0	762	5	BQ121182
40	188.2	15.8	882	7	CN125340
41	186.6	15.6	652	6	CA255028
42	184.6	15.5	450	8	BH509846
43	183.8	15.4	733	6	CB292744
44	183.6	15.4	830	8	BZ500536
C 45	182.4	15.3	735	4	BM408950

ALIGNMENTS

RESULT 1
CNS0A29J
LOCUS
DEFINITION
CNS0A29J 1340 bp mRNA linear HTC 06-FBB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTSL712A05 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION
BX827536
VERSION
GI:42459416
KEYWORDS
HTC; GSLT cDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 1340)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation

JOURNAL
REFERENCE
2 (bases 1 to 1340)
Unpublished
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen)
full-length libraries construction : Temple G
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
Location/Qualifiers
1. 1340
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"

FEATURES
source

Query Match		91.4%;	Score 1091.2;	DB 3;	Length 1359;
Best Local Similarity		97.4%;	Pred. No. 2.1e-301;		
Matches 1130;		Conservative	0;	Mismatches 28;	Indels 2; Gaps 2;
QY	37	GGTGACCAAGCTTGAAGCAAACTTATAGATCATGAGGTGGTAACTCAATCATCATCA	96		
DB	39	GGTGACCAAGCTTGAAGCAAACTTATAGATCATGAGGTGGTAACTCAATCATCATCA	98		
QY	97	TCAGCTGTGCTCAAAACCCAGAGACTATAAAAGGTGGCTTCGTGTCTCCATATACGTAATC	156		
DB	99	TCAGCTGTGCTCAAAACCCAGAGACTATAAAAGGTGGCTTCGTGTCTCCATATACGTAATC	158		
QY	157	TTTGTCCTCTTTTGCAGGCCATAGCTACAAATTCCTGGGTAGATTGTACTATGAATGA	216		
DB	159	TTTGTCCTCTTTTGCAGGCCATAGCTACAAATTCCTGGGTAGATTGTACTATGAATGA	218		
QY	217	GGAAATAGACATATGTGTACACATCTTCAACTCATTTGGCTCCCTGACTGTTCTCG	276		
DB	219	GGAAATAGACATATGTGTAAACATCTTCAACTCATTTGGCTTCCTGACTGTTCTCG	278		
QY	277	TTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCC	336		
DB	279	TTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTCACTCAGTCC	338		
QY	337	CTTCTCTTCAACACCTTGTGATCGGTTTACTGTGCACTGGACTGTAGTTCGCTTAT	396		
DB	339	CTTCTCTTCAACACCTTGTGATCGGTTTACTGTGCACTGGACTGTAGTTCGCTTAT	398		
QY	397	GCTTATTTCTGTCAGTGGTGTCTACTTACCAGTCTCTACTTTTCTCCCTCATCTTG	456		
DB	399	GCTTATTTCTGTCAGTGGTGTCTACTTACCAGTCTCTACTTTTCTCCCTCATCTTG	458		
QY	457	GCTCACAGTTGGCTTCTACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT	516		
DB	459	GCTCACAGTTGGCTTCTACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT	518		
QY	517	CTTTGATAGTCACTTTTCTCTCTCACTGATCTCTGCTCTTCTTGTGTCAC	576		
DB	519	CTTTGATAGTCACTTTTCTCTCTCACTGATCTCTGCTCTTCTTGTGTCAC	578		
QY	577	ACTGATTCAGAACTCAACTATGTCTAGAGTACATGTGATCGGTTCTATGT	636		
DB	579	ACTGATTCAGAACTCAACTATGTCTAGAGTACATGTGATCGGTTCTATGT	638		
QY	637	ACCATCGGTCTTCCCTGGGATGGACTGTACTATCTCTGATACAAATGTCTTCAGG	696		
DB	639	ACCATCGGTCTTCCCTGGGATGGACTGTACTATCTCTGATACAAATGTCTTCAGG	698		
QY	697	AAAGTTTTCAGGAAGCATATCTCAGCAGTCAAGACTTGGCCATTTACCAGTCTCTA	756		
DB	699	AAAGTTTTCAGGAAGCATATCTCAGCAGTCAAGACTTGGCCATTTACCAGTCTCTA	758		
QY	757	GTTCGAGTGTGTAGTTCTCATAGACATTTTTCAGGTGGAGTGGGAACTTTGCCA	816		
DB	759	GTTCGAGTGTGTAGTTCTCATAGACATTTTTCAGGTGGAGTGGGAACTTTGCCA	818		
QY	817	AGTGAGATGAGAACTACAAACTCGGGAAGTGTCTATGTTTGTACTTTTACCTCGGCA	876		
DB	819	CGTGAGATGAGAACTACAAACTCGGGAAGTGTCTATGTTTGTACTTTTACCTCGGCA	878		
QY	877	GCTATTTCCTGGCAAGTCTACACTCTCGTCTTGTGGGAT-TGATCTTCG-AGTCATCT	934		
DB	879	ACTATTTCATGGCAAGTCTACACTCTTGTGTGTGGGATGATCTTCGCACTCATCT	938		
QY	935	CTGTGTTCTTCCAAATCCAACTGTGGGATTTGCCATAGTTCCAGTTGGCGGAGTGA	994		
DB	939	CTGTGTTTCCAAATCCAACTGTGGGATTTGCCATAGTTCCAGTTGGCGGAGTGA	998		
QY	995	TAGTTTTTCCATGATGAATGGAGCATCCAAATCTTCTCCATTTATTTAGTCTATCTGG	1054		
DB	999	TCGTTTTTCCATGATGAATGGAGCATCAAAATCTTCTCCATCATTTTAGTCTATCTGG	1058		
QY	1055	GCTTCTCTTCAITTCGTCTATCAGCACTACCTCGACGAAAGAGTTGAATACTAGCCACA	1114		

Db	1059																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
----	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db	243	GGGAAAGCACATATGTGTGAACACTTCTTCAACTCAITGGGCTTCCTCTACTGATCTC	302
Qy	277	TTCCGCTCTTTTCTCGAATCAGGCAACCAATCAACAGATACAAATTTCACTCAGTCC	336
Db	303	TTCCGCTCTTTTCTCGAATCAGGCAACCAATCAACAGATACAAATTTCACTCAGTCC	362
Qy	337	CTTCTCTTCCACCACTTGTGATCGGTTACTTGTGACATGGAGTGTAGTGTCCGTTAT	396
Db	363	CTTCTCTTCCACCACTTGTGATCGGTTACTTGTGACATGGAGTGTAGTGTCTGTTAT	422
Qy	397	GCTTATTTCTGTGACATGGTGTCTTACTTACCACTCTCTACTTTCTCCCTCATCTTG	456
Db	423	GCTTATTTCTGTGACATGGTGTCTTACTTACCACTCTCTACTTTCTCCCTCATCTTG	482
Qy	457	GCTCAGTGTGCGCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT	516
Db	483	GCTCAGTGTGCGCTTCACTGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT	542
Qy	517	CTTTGATAGTGTCTTGTCTTCTCTCACTGTATTCCTCTGCTCTTCTTGTGTCAAC	576
Db	543	CTTTGATAGTGTCTTGTCTTCTCTCACTGTATTCCTCTGCTCTTCTTGTGTCAAC	602
Qy	577	ACTGATTCAGAAACTCAACTAATGTATCTAGAGTACAGTATGTATCGGTTTCTATGT	636
Db	603	ACTGATTCAGAAACTCAACTAATGTATCTAGAGTACAGTATGTATCGGTTTCTATGT	662
Qy	637	ACCATCGGCTTCCGCTGGGATTTGAGCTTTACTATCTCTGATACAAATGCTTTCAGG	696
Db	663	ACCATCGGCTTCCGCTGGGATTTGAGCTTTACTATCTCTGATACAAATGCTTTCAGG	722
Qy	697	AAAGTTTCCAGAGCATACATCTCAGCAGTCAGGACTTGGGCAATTTACCACTCTCTA	756
Db	723	AAAGTTTCCAGAGCATACATCTCAGCAGTCCTGTACTTGTCCAAATTTACCACTCTCTA	782
Qy	757	GTTCGAGTGTGTAGTCTCTATAGACTTTTTCGAAGTGGAGTGGGAACTTTGGCA	816
Db	783	GTTCGAACTTTGTGTGATCTCATAGGACTGTTTTCGAAGTGGAGTGGGAACTTTGGCA	842
Qy	817	AGTGAGATGAGAACTACAAACTCGGGAAGTGTATATGTTTTCGATTTTACCTCGGCA	876
Db	843	AGTGAGATGAGAACTACAAACTCGGGAAGTGTATATGTTTTCGATTTTACCTCGGCA	902
Qy	877	GCTATTTCTCGCAAGTCTACACTCTCTGTCTTGTGGGATGATCTTTCGAGTCATCTCT	936
Db	903	GCTATTTCTCGCAAGTCTACACTCTGTGTGTGTGGGATTAATCTTCAGTCTTCTCT	962
Qy	937	GTGTTCTCCAATTCATACAGCTGGGATTTGCCTATAGTTCAGTTCGCGAGTGATA	996
Db	963	GTGTTCTCCAATTCATACAGCTGGGATTTGCCTATAGTTCAGTTCGAGTGATA	1022
Qy	997	GTGTTCTCCAATTCATACAGCTGGGATTTGCCTATAGTTCAGTTCGCGAGTGATA	1056
Db	1023	GTGTTCTCCAATTCATACAGCTGGGATTTGCCTATAGTTCAGTTCGCGAGTGATA	1082
Qy	1057	TTCTCTTTCATTCGTCTATCAGCACTACCTTCGACGAAAGAGTGAATCTAGCCACACA	1116
Db	1083	TTCTCTTTCATTCGTCTATCAGCACTACCTTCGACGAAAGAGTGAATCTAGCCACACA	1142
Qy	1117	AGTGTGTAGGAGA 1130	
Db	1143	AAACCTGTTGAGGA 1156	
RESULT 5			
CN50A2U8			
LOCUS			
DEFINITION			
Arabidopsis thaliana Full-length cDNA Complete sequence from clone			
GSLTSL592F07 of Adult vegetative tissue of strain col-0 of			
Arabidopsis thaliana (thale cress).			
BX827470			
ACCESSION			
BX827470.1 GI:42460504			
VERSION			
KEYWORDS			
HTC; GSLT_cDNA.			
SOURCE			
Arabidopsis thaliana (thale cress)			

ORGANISM	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
REFERENCE	1 (bases 1 to 1312)			
AUTHORS	Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.			
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 1312)			
AUTHORS	Genoscope.			
TITLE	Direct Submission			
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)			
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis Genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full-length http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.			
FEATURES	Location/Qualifiers			
source	1..1312			
	/organism="Arabidopsis thaliana"			
	/mol_type="mRNA"			
	/strain="Col-0"			
	/db_xref="taxon:3702"			
	/clone="GSLTSL592F07"			
	/tissue type="Adult vegetative tissue"			
	/plasmid="pCMVSPORT_6"			
gene	1..1312			
	/gene="At4g18205"			
ORIGIN				
Query Match	75.6%;	Score 903;	DB 3;	Length 1312;
Best Local Similarity	90.4%;	Pred. No. 2.1e-247;		
Matches	990;	Conservative 0;	Mismatches 95;	Indels 10; Gaps 2;
Qy	37	GGTGACCAAGAACTTAGAAGCAAACTTATAGATCATGAGTGGTAACATCATCA	96	
Db	86	GGTGACCAAGAACTTAGAAGCAAACTTCTAGATCATGAG-----GAAACTGAATCA	136	
Qy	97	TCAGCTGTGCTTCAACCGAGAACTATAAAGGTGGCTTCGTCTCTCATATACGTATTC	156	
Db	137	TTTTTCAGTACCTCAGACGAAGAACTGTAAGAGTGGCTCCGTGTCTCCATATACGCAATC	196	
Qy	157	TTTGTCTCTTTTTCGCAGCCACTAGTACAATCTTGGGTAGATTGTACTATGAAATCGA	216	
Db	197	TTTGTCTCTTCTGCCAACCACTTGTCTACAGTCTTGGGTAGACTGTACTATGAGAATGGA	256	
Qy	217	GGAAATAGCACATATGTGTAACACTTCTTCAACTCATTTGGCTTTCCCTGTACTGGTTCTG	276	
Db	257	GGGAAAGCACATATGTGTTAACTTCTTCAACTCATTTGGCTTTCCCTGTACTGATTCTC	316	
Qy	277	TTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC	336	
Db	317	TTCCGCTTCTTTTCTCGAATCAGGCAACCCAAATCAACAGATACAAATTTTCAGTCAGTCC	376	
Qy	337	CCTTCTCTTCCACCACTTGTGATCGGTTTACTTGTGCACTGGAGTGTAGTGTCCGCTTAT	396	
Db	377	CCTTCTCTTCCACCACTTGTGATCGGTTTACTTGTGCACTGGAGTGTAGTGTCTGCTTAT	436	
Qy	397	GCTTATTTGTGTCAGTAGGGTTGCTCTACTTACCAGTCTCTACTTTTCTCCCTCATCTTG	456	

Db 437 GCTATTGTCGACAGTTGGTGGTTCCTTACCTACAGTCTCTACTTCTCCTCATCTTG 496
Qy 457 GCCTCACAGTTGGCCTTCACAGCTTTTCTCAATATTTCTTAACTCGCAAAAGTTCACT 516
Db 497 GCCTCACAGTTGGCCTTCACAGCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACT 556
Qy 517 CTTTGTAGTACAGTTCTTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 576
Db 557 CTTTGTAGTACAGTTCTTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 616
Qy 577 ACTGATTCAGAAAACTCAACTAATGTATCTAGAGTACAGTATGTGA-TCGGGTTTCATATG 635
Db 617 ACTGATTCAGAAAACTCAACTAATGTATCTAGAGTACAGTATGTGA-TCGGGTTTCATCTG 676
Qy 636 TACCATCGGTGCTTCGGCTGGGATTTGGAGTCTTCTCTCTCTCTCTCTCTCTCTCTCT 695
Db 677 TACCATCGGTGCTTCGGCTGGGATTTGGAGTCTTCTCTCTCTCTCTCTCTCTCTCTCT 736
Qy 696 GAAAGTTTTCAGGAGCATACATCTCTCAGCAGTCAAGCAGTTGGCCATTTACCAAGTCTCT 755
Db 737 GAAAGTTTTCAGGAGCATACATCTCTCAGCAGTCAAGCAGTTGGCCATTTACCAAGTCTCT 796
Qy 756 AGTTGCGAGTTGTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGTGGGAAACTTTGGCC 815
Db 797 AGTTGCAACTGTGTGGTACTCATAGGACTGTTCAGAGTGGAGTGGGAACTCTCGCC 856
Qy 816 AAGTGAGATGAGAACTCAAACTCGGAAAGTGTCTATATTTTGTAGCTTTAGCCTCGGC 875
Db 857 AAGTGAGATGAGAACTCAAACTCGGAAAGTGTCTATATATCTTGACTTTGGCCTCAGC 916
Qy 876 AGCTATTTCTCGGCAAGTCTACATCTCTGGTCTTGTGGGATGTCTTCGAGTCTCTCT 935
Db 917 AGCTATTTCTCGGCAAGTCTACATCTCTGGTCTTGTGGGATTAATCTTCGAGTCTCTCT 976
Qy 936 TGTGTTCTCCAAATCCATAACAGCTGTGGGATTTGCTATAGTTCAGTTCGCGGAGTGAT 995
Db 977 TGTGTTCTCCAAATCCATAACAGCTGTGGGATTTGCTATAGTTCAGTTCGAGTGTAG 1036
Qy 996 AGTTTTCATGATGAGATGAGCAGTCCAAATCTTCCATATATTTTGTAGTCTATCGGG 1055
Db 1037 AGTTTTCATGATGAGATGAGCAGTCCAAATCTTCCATATATTTTGTAGTCTATCGGG 1096
Qy 1056 CTTCTTTCTATCGTCTATCGACTACCTCGAGGAAAGTGTGATATCTAGCCACAC 1115
Db 1097 CTTCTTTCTATGTTATCGACTACCTCGAGGAAAGTGTGATATCTAGCCACAC 1156
Qy 1116 AAGTCTCTAGGAGA 1130
Db 1157 AAACCTGTTGAGGA 1171

RESULT 6
B77316/c
LOCUS T3213TP TAMU Arabidopsis thaliana genomic clone T3213, genomic 699 bp DNA linear GSS 16-JAN-1998
DEFINITION survey sequence.
B77316
VERSION B77316.1 GI:2773955
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 699)
REFERENCE Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K.,
AUTHORS Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and
Venter,J.C.
A BAC End Sequence Database for Identifying Minimal Overlaps in
Arabidopsis Genomic Sequencing. Update 3
JOURNAL Unpublished (1997)
COMMENT Contact: Steve Rounsley

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13-21
Class: BAC ends
High quality sequence stop: 699.
FEATURES
Location/Qualifiers
source 1..699
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="T3213"
/sex="hermaphrodite"
/clone_lib="TAMU"
/notes="Vector: BelobACII; Site_1: HindIII; Site_2:
HindIII; Produced by Rod Wing"

ORIGIN
Query Match 41.1%; Score 491; DB 8; Length 699;
Best Local Similarity 88.4%; Pred. No. 2.9e-129;
Matches 533; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
Qy 528 CAGTCTTCTGCTCTCTCACTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 587
Db 699 CAATCTTTTATCTCTCTACGGTTTCTCTGCCCCCTCTCTCTCTCTCTCTCTCTCTCT 640
Qy 588 AAATCAACTAATGTATCTAGAGTACAGTATGTATCGGGTTCATATGTACCATCGGTGC 647
Db 639 AAACACAAATATCTAGAGTACAAATATGATTTGGGTTCATCTGTACCATGTTGTTGC 580
Qy 648 TTCCGCTGGGATTTGAGTCTTACTCTCTGATACAAATCTCTTTCAGGAAAGTTTTCAC 707
Db 579 TTCCGCTGGGATTTGAGTCTTACTCTCTCTGATACAACTCTCTTTCAGGAAAGTTTTCAC 520
Qy 708 GAGCATACATCTCTCAGCAGTACGAGCTTTGGCCATTTACAGTCTCTAGTTGGAGTTG 767
Db 519 GAAGCATACATCTCTCAGCAGTCTCTGGACTTTGGCCATTTACCACTCTCTAGTTGCA 460
Qy 768 TGTAGTTCTCATAGGACTTTTTCAGAGTGGAGAGTGGGAACTTTTGGCAAGTGAGATGAG 827
Db 459 TGTGTTACTCATAGGACTTTTTCAGAGTGGAGAGTGGAGAACTCTGCAAGTGAGATGAG 400
Qy 828 AAATACAAACTCGGAAAGTGTCTATGTTTGTGACTTTTACCTCTCGCAGCTATTTCTCTG 887
Db 399 AAATACAAACTCGGAAAGTGTCTATATCTTGTGACTTTTGGCCCTCAGCAGCTATTTCTG 340
Qy 888 GCAAGTCTACACTCTCTGCTTGTGGGATTTGATCTTCGAGTCTCTCTCTCTCTCTCTCAA 947
Db 339 GCAAGTATACACTCTTGTGTTGTGGGATTAATCTTCAGTCTCTCTCTCTCTCTCTCAA 280
Qy 948 TTCCATAACAGCTGTGGGATTTGCTATAGTTCAGTTTGGCGAGTGTAGTTTTCATGA 1007
Db 279 TTCCATAACAGCTGTGGGACTGCTCTATAGTTCAGTTTGTAGCAGTGTAGTTTTCATGA 220
Qy 1008 TAGAATGAGCAGTCCAAAATCTTCTCATATTTTGTAGCTATCTCGGGCTCTCTCTCT 1067
Db 219 TAAGTGGATGCATCAAAAGATTTTCTCCATCAATTTTAGCTATCTCGGGCTCTCTCTCT 160
Qy 1068 CGTCTATCAGCACTACCTCCAGCAAGAGTTGAAATCTAGCCACACAGTGTGTAGG 1127
Db 159 TGTCTATCAGCACTACCTCCAGCAAGAGTTGAAAGCTTGGCAGACAAACCTGTTGA 100
Qy 1128 AGA 1130
Db 99 GGA 97

RESULT 7
BH498028

LOCUS BH498028 709 bp DNA linear GSS 13-DEC-2001
DEFINITION BOHQ68TF BOHQ Brassica oleracea genomic clone BOHQ68, genomic survey sequence.
ACCESSION BH498028
VERSION BH498028.1 GI:17706132
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eursoids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 709)
REFERENCE Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
AUTHORS Whole genome shotgun sequencing of Brassica oleracea
TITLE Unpublished (2001)
JOURNAL
COMMENT Other_GSSs: BOHQ68TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.

```

FEATURES
    source
        Class: sheared ends.
        Location/Qualifiers
            1..709
                /organism="Brassica oleracea"
                /mol_type="genomic DNA"
                /strain="Tol000DH3"
                /db_xref="taxon:3712"
                /clone="BOHQK68"
                /clone_lib="BOHQ"
                /note="Vector: pHS01; Site 1: BstXI; 2-3 kb sheared
                genomic DNA inserted into pHS01 using BstXI linkers"
ORIGIN

```

Query Match	37.1%;	Score 442.6;	DB 8;	Length 709;
Best Local Similarity	81.2%;	Pred. No. 2.2e-115;		
Matches 556;	Conservative 0;	Mismatches 114;	Indels 15;	Gaps 3;
Qy	1	TCATGAGATATAAATACATGAGTGTAAATTTTTTCAGGTGACCGAGAACTTTAGAGCAAAC	60	
Db	37	TCATGAATTCGTGAATCATGAGTGTAAATTTTTTCAGGTGACCGAGAACTTTAGAGCAAAC	96	
Qy	61	CTTATAGATCATGAGGTGGTAACTGAAATCATCATCATCAGCTGTGCCTCAAAACCGAGAAC	120	
Db	97	CTGATAGAT-----GCACTGAACCAATCATCATCA---GTACCTCAAAACGAAGAC	144	
Qy	121	TATAAAGGTGGCTTCGTGCTCCATATACGTAATCTTTGTCTCTTTTCCAGCCACTA	180	
Db	145	TATAAAGATGGATCCGTCTCTCCATCTATGTCTTCTTCGTCTCTTCTTCCAAACCATT	204	
Qy	181	GCTACAAATTCGGGTAGATTGTACTATGAAAAATGGAGGAAATAGACATATGTGGTAAACA	240	
Db	205	GCCACAAATTCGGGTAGACTGTACTAGAAAAACGGTGGAGAGAGCATACGTGGTAAACG	264	
Qy	241	CTTCTTCAACTCATTTGGCTTCCCTGTACTGGTCTGTCTCGCTCTCTTTCTCGAATCAG-	299	
Db	265	CTTCTCCCACTCAFTGGCTTCCCTGTTTGTATCCTCTTCAACTTCTTTTTTCACAACTCAGA	324	
Qy	300	--GCAACCCCAATCAACAGATACAAATTTTCAGTCAGTCCCTCTCTTCAACACCCCTTGCA	357	
Db	325	CAACACCAAAATCAACAGATACAAATGTCAACGAGTCCCTCTCTTCAACACCCCTTTTA	384	
Qy	358	TCGGTTTACTTGTGCACTGGACTGCTAGTGTCCGCTTATGTCTTATTTGTCTGCAGTAGGG	417	
Db	385	TTAGTTTACATGTGCACCGGACTTATACTCTCTGCTTACGCTTATCTCTATGCAATTGGT	444	
Qy	418	TTGCTCTACTTACAGACTCTCTACTTTTCTCCCTCATCTTTGGGCTCAACAGTTGGCGTTCACT	477	
Db	445	TTACTCTACTACTACTGTTTCCACCTTCTCCCTCATCTTTGGCCCTCAACAGCTTAGCGTTCACT	504	

Qy	478	GCCTTTTCTCATATTTCTTAACTCGCAAAAGTTCACCTCTTTGTAGATGATCAGTTCTTTG	537		
Db	505	GCCTTCTTCTTCTTACTTCTTAACTGCGCAAAAGTTCACCTCCCTTCATAGTCAACTCTCTG	564		
Qy	538	CTTCTCCTCACTGTCATCCTCTGCTCTCTTGTGGTCAACACTGATTTTCAGAAACTCAACT	597		
Db	565	TTTTTCTCTTACCCTGCTCTCTGCGCTCTCTCGTGGTTGACACCGAGTCAACAAGACACA	624		
Qy	598	AATGTATCTAGAGTACAGTATGTGATCGGGTTCATATGTATGATACCATCGGTGCTTCGGCTGGG	657		
Db	625	AATGCTCTAGACTACAGTATGTGATCGGGGTATATGTCACCATCCGTGCTTCAGCTGGG	684		
Qy	658	ATTGGACTGTTACTATCTCTGATAC	682		
Db	685	ATAGGACTGTGCTATCTCTGATAC	709		
RESULT	8				
LOCUS	BZ087060	691 bp	linear		
DEFINITION	lkh10d12.b1 B.oleracea002 Brassica oleracea genomic, genomic survey sequence.				
ACCESSION	BZ087060				
VERSION	BZ087060.1	GI:23719662			
KEYWORDS	GSS.				
SOURCE	Brassica oleracea				
ORGANISM	Brassica oleracea				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.				
AUTHORS	1 (bases 1 to 691) Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T., Nash, W., Rabinowicz, P.D. and Wilson, R.K.				
TITLE	Whole genome shotgun reads from Brassica oleracea				
JOURNAL	Unpublished (2002)				
COMMENT	Contact: Richard K. Wilson Genome Sequencing Center Washington University School of Medicine Email: submission@watson.wustl.edu Plate: lkh10 row: d column: 12 Seq primer: -21UPpOT forward Class: shotgun High quality sequence start: 18 High quality sequence stop: 491.				
FEATURES	Location/Qualifiers				
source	1..691 /organism="Brassica oleracea" /mol_type="genomic DNA" /db_xref="taxon:3712" /clone_lib="B.oleracea002" /notes="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea TO1000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."				
ORIGIN					
Query Match	36.6%	Score 437.4;	DB 8; Length 691;		
Best Local Similarity	80.4%;	Pred. No. 6.7e-114;			
Matches	525; Conservative	0; Mismatches 126;	Indels 2; Gaps 1;		
Qy	393	TTATGCTTATTTCTGTCAGTAGGGTTGCTCTACTTACCAGTCTCTACTTTCTTCTCCCTCAT	452		
Db	33	TTCCGCTTAFTCTATGCAATGGTTTACTCTACTACCTGTTTCCACCTTCTCCCTCAT	92		
Qy	453	CTTGGGCTCACAGTGGCCCTTCACTGCCTTTTCTCATATTTCTTAACCTCGCAAAAGTT	512		
Db	93	CTTGGGCTCACACTAGCCTTCACTGCCCTTCTCTTACTTCTTAACTCGCAAAAGTT	152		
Qy	513	CACCTCCTTTGATAGTCAGTTCTTTTGCTTCTCCTCACTGATCCTCTCTCTTCTTGCT	572		

```
Db 153 CACTCCCTTCATAGTCAACTCTCTGTTTCTCCTTACCGTCTCCTCTGCCCTCCCTCGTGGT 212
Qy 573 CAACATGATTCAGAAAACTCAACTAATGTATCTAGAGTACAGATGTGATCGGGTTCAAT 632
Db 213 TGACACCGAGTCAACAAGACACAAATATGTCTCTAGACTACAGTATGTGATCGGGTTAT 272
Qy 633 ATGTACCATCGGTGCTTCGGTGGGATTTGGAGTGTACTATCTCTGATACAAATGCTCTT 692
Db 273 ATGCACCATCGGTGCTTCAGCTGGGATAGGATGCTGCTATCTCTGATACAACTCCCTCTT 332
Qy 693 CAGGAAGCTTTTCAGAAAGCATACATCTCAGCAGTACAGGACTTGGCCATTTTACCAGTC 752
Db 333 CAGGAAGCTTTTCAGAGAGCATACATCTCAGTGTATGACTTGTACTATCTTACCAGTC 392
Qy 753 TCTAGTTCGAGTGTGTAGTTCATAGGACTTTTTCGAAAGTGAGAGTGGGAACTTT 812
Db 393 TCTAGTTCGAAAGTGCCTGCTCTCATTTGAGACTTTTTCGTGAGCGGAGAGTGGAGAACTTT 452
Qy 813 GCCAGTCAGATGAGAACTCAACTCGGAAAGTGTATATGCTTTTGTAGCTTTAGCCTC 872
Db 453 GCCTAGTGAAGATGAGAACTCAAACTTGGGCAAGTGTCTTATGTTGTGAGCTTGGICTC 512
Qy 873 GGCAGTATTTCTCGGCAAGTCTACACTCTCTGGTCTTGTGGGATTTGATCTTCGAGTCATC 932
Db 513 TGCAGTATCTCTTGGCAAGTCTACACTGTGTGGCTTGGGGGTTGAACTTTGATCGTC 572
Qy 933 CTCTGTGTTCTCCAAATCCATAACAGCTGTGGGATTTGCCATAGTTCAG--TTGCCGCA 990
Db 573 CTCTGGGCTCTCCAACTCCATAACCGCGTGGGGTGCCTATATAACGGGCCATAAGCT 632
Qy 991 GTGATAGTTTCCATGATGAGATGAGCAGCATCCAAATCTTCTCCATTTT 1043
Db 633 GGATAATTTTCCATGAAAAAATGGAAGCGTCCAAATCTTCTTCCATCGTTT 685
```

```
RESULT 9
AV828990 LOCUS AV828990 RAF9 Arabidopsis thaliana cdna clone RAFL09-36-N15 5',
DEFINITION mRNA sequence.
ACCESSION AV828990
VERSION AV828990.1 GI:19871050
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 558)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cdna (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekio@tc.riken.go.jp
An Arabidopsis full-length cdna library was constructed essentially
as reported previously (Seki et al., 1998). cdna cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further
details.
```

```
FEATURES
source Location/Qualifiers
1..558
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
```

```
/db_xref="taxon:3702"
/clone="RAFL09-36-N15"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAFL9"
/notes="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"
ORIGIN
Query Match 32.3%; Score 386; DB 1; Length 558;
Best Local Similarity 91.0%; Pred. No. 3.5e-99;
Matches 423; Conservative 0; Mismatches 33; Indels 9; Gaps 1;
Qy 37 GGTGACCAGAACTTTAGAACCAACCTTATAGATCATGAGGTGGTAACGTGAATCATCATCA 96
Db 103 GGTGACCAGAACTTTAGAACCAACCTTATAGATCATGAG-----GAAACTGGAATCA 153
Qy 97 TCAGCTGTGCTCAACCGGAGAACTATAAAGGTGGCTTCGTCTCCATATAGCTAATC 156
Db 154 TTTTCAGTACCTCAACGAGAACTGTAGAGGTGGCTCCGTCTCCATATACGCAATC 213
Qy 157 TTTGTCTCTTTTCCAGGCACTAGCTACAAATCTGGGTAGATTGTACTGAAAAATGGA 216
Db 214 TTTGTCTCTTTTCCAGGCACTAGCTACAAATCTGGGTAGATTGTACTGAAAAATGGA 273
Qy 217 GGAATAGCATATGTGGTAAACATCTTCAACTCAATCGGTTCCCTGTACTGGTTCTG 276
Db 274 GGGAAAAGCATATGTGGTAAACATCTTCAACTCAATCGGTTCCCTGTACTGATTTCTC 333
Qy 277 TTCCGCTCTTTTCTCGAATCAGGCAACCAATCAACAGATACAAATTTTCAGTCAGTCC 336
Db 334 TTCCGCTCTTTTCTCGAATCAGGCAACCAATCAACAGATACAAATTTTCAGTCAGTCC 393
Qy 337 CCTTCCTTCAACCCCTTGCATCGTTTACTTGTGCACTGGACTGTAGTGTCCGCTTAT 396
Db 394 CCTTCCTTCAACCCCTTGCATCGTTTACTTGTGCACTGGACTGTAGTGTCCGCTTAT 453
Qy 397 GCTTATTTGTCTGAGTGGTGTCTCTACTTACCAGTCTCTACTTTTCCCTCATTTG 456
Db 454 GCTTATTTGTCTGAGTGGTGTCTCTACTTACCAGTCTCTACTTTTCCCTCATTTG 513
Qy 457 GCCTCAGATTTGGCTTCACTGCTTTTCTCATATTTTCTTAAAC 501
Db 514 GCCTCAGATTTGGCTTCACTGCTTTTCTCATATTTTCTTAAAC 558
```

```
RESULT 10
AI996743/c LOCUS AI996743 446 bp mRNA linear EST 08-SEP-1999
DEFINITION 701668184 A. thaliana, Columbia Col-0, root-1 Arabidopsis thaliana
cdna clone 701668184, mRNA sequence.
ACCESSION AI996743
VERSION AI996743.1 GI:5843648
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 446)
Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D.,
Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P.,
Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D.,
Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C.,
Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S.,
Nobriga, A., Murry, B., Turner, C., Krikorian, S., Elder, L. and
Hanson, D.
```

```
TITLE Arabidopsis thaliana Gene Expression MicroArray
JOURNAL Unpublished (1999)
COMMENT Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
```

Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.

FEATURES

source
1. .446
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="701668184"
/tissue_type="root"
/dev_stage="4 - 7 weeks"
/clone_lib="A. thaliana, Columbia Col-0, root-1"
/note="Vector: pSPORT; Site 1: NotI; Site 2: SalI; cDNA library was derived from untreated root tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunted, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

ORIGIN

Query Match 28.0%; Score 334.8; DB 1; Length 446;
Best Local Similarity 99.1%; Pred. No. 1.6e-84;
Matches 336; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 856 GTTTGACCTTTAGCCCTCGGCAGCTATTTCCTGCGAAGTCTACACTCTCGTCTGTGGGA 915
DB 446 GTTTTGACCTTTAGCCTCGGCAGCTATTTCCTGCGAAGTCTACACTCTCGTCTGTGGGA 387

QY 916 TTGATCTTCAGTCATCCTCTGTGTTCTCAATTCATTAACAGCTGTGGGATTGCGCTATA 975
DB 386 TTGATCTTCAGTCATCCTCTGTGTTCTCAATTCATTAACAGCTGTGGGATTGCGCTATA 327

QY 976 GTTCCAGTTGCGGCAGTGTAGTTTCCATGATAGATGCGCGCATCCAAATCTTCTCC 1035
DB 326 GTTCCAGTTGCGGCAGTGTAGTTTCCATGATAGATGCGCGCATCCAAATCTTCTCC 267

QY 1036 ATTATTTTACGTCATCGCGCTTCTTTTCATTCGTCATCAGCACTACCTCGACGAAAG 1095
DB 266 ATTATTTTACGTCATCGCGCTTCTTTTCATTCGTCATCAGCACTACCTCGACGAAAG 207

QY 1096 AAGTTGAATACAGCACACAAGTGTGTAGGAGATCTTCATCTACCTGTGAGGAAGT 1155
DB 206 AAGTTGAATACAGCACACAAGTGTGTAGGAGATCTTCATCTACCTGTGAGGAAGT 147

QY 1156 CACACAAACATCAAAAGTGTGTATCAAAAGCATATTTC 1194
DB 146 CACACAAACATCAAAAGTGTGTATCAAAAGCATATTTC 108

RESULT 11

CNS0A3B5 1519 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFB632D12 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).
BX826795
BX826795.1 GI:42460970
HTC; GSI: cDNA.
DEFINITION Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1519)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,

TITLE

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 1519)

AUTHORS

Genoscope.

TITLE

Direct Submission

JOURNAL

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT

- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length
http://www.genoscope.cns.fr/cgi-bin/ggb?ggb?source=Arabidopsis.

FEATURES

source

Location/Qualifiers

1. .1519

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="Col-0"

/db_xref="taxon:3702"

/clone="GSLTFB632D12"

/tissue_type="Flowers and buds"

/plasmid="pCMVSPORT_6"

1. .1519

/gene="At4g18210"

ORIGIN

Query Match 26.5%; Score 317; DB 3; Length 1519;
Best Local Similarity 59.7%; Pred. No. 3.1e-79;
Matches 641; Conservative 0; Mismatches 420; Indels 12; Gaps 6;

QY 96 ATCAGCTGTGCTCAAAACGAGAACTATAAAGGTGGCTTCGTGCTCCATATACGTA-A 154
DB 176 AGCAAGATATCTTACTCTAACACATACAAACGGTGGCTCAGGGTGACTTATATACATT 235

QY 155 TCTTTGCTCTTTTGGCCAGCCACTAGCTACAAATCTCGGTAGATGTGACTATGAATG 214
DB 236 TCTTTGCTCATTTAGGCCCAACAGTTGCTACAAATTTGGCGAGATATGATGACAG 295

QY 215 GAGGAATAGCACAT-ATGTGTTAACTCTTCAACTCATTTGGCTTCCCTGTACTGGTT 273
DB 296 GAGGAACAGTAATGGCTTAGCAACGGTAGTTCAACTTTGTTGGCTTCTGTTTACTT 355

QY 274 CTGTTCCGCTCTTTTCTCGAATCAGGAACCCAAATCAACAGATACAAATTTTACGTAG 333
DB 356 CCATATTATATCTTGTCTATTTA-----AAACACATGCAACAACTGATAGATGAAAA 409

QY 334 TCCCTTCTTCAACACCTTGCATCGGTTTACTTGTGCACTGGACTGCTAGTGTCCGCT 393
DB 410 AGAACCTCACCCTAGGAACCGTATGTTGTTTACGTAAGTGTGGACTTCTTGTAGGACA 469

QY 394 TATGCTTATTTGCTGCAAGTGGGTTGCTCTACTTTACCAGTCTCTACTTTTCTCCCTCATC 453
DB 470 GATTGCTATTTGTAATCCATTTGGACTTCTTACTTACCGGTTTCTACCTATTTCCCTGATC 529

QY 454 TTGGCCTCAGATTGGCCTTCACTGCCTTTTCTCATATTTCTTAACTCGCAAAAGTTC 513
DB 530 TGTGCTCTCAGTTAGCCTTCAATGCTTTTCTTCTTATTTTCTTAACTCAAAAAAATT 589

QY 514 ACTGCTTTGATAGTCAAGTTCTTTGCTTCTCTCACTGTATCTCTGCTCTTCTTCTGTC 573
DB 590 ACCCTATCATTTAAATTTCTTTTCTCTCTAACTATATTTTCCACCCCTACTTGCATTC 649

```
Qy 574 AACACTGATTCAGAAACCTCACTAATGTAATCTAGAGTACAGTATGTGATCGGGTTTCATA 633
|||
Db 650 AATAATGAGGACAGATTCACAAAAGTTACAAAAGGAGATTTTTTAAAGGTTTCATA 709
|||
Qy 634 TGTACATCGGTGCTTCGGTGGGATGAGCTGTTACTATCTCTGATACA-AAATGCTCTT 692
|||
Db 710 TGCACCGTTGCTGCTTTGCTGGTTATGTTTATTTATCCCTACAAACAGCAGCGCTTT 769
|||
Qy 693 CAGGAAAGTTTTCAGAAAGCATACATCTCAGCAGTACGAGCTGCGCAATTTACCAAGTC 752
|||
Db 770 TTTAAAGTCTTAAAGAGCAAAATTTCTCAGAAGTTATGGATATGATAATCTCATAGAA 829
|||
Qy 753 TCTAGTTCGAGTGTGTAGTTCTCATAGGACTTTTTGCAAGTGGAGAGTGGGAAACTTTT 812
|||
Db 830 TCTGATGGGAGTTGTGTTCCGGTGGTGGGCTTTTCTCACAATGAATGGAATAATTT 889
|||
Qy 813 GCCAAGTGAAGTGAAGAACTCAAACTCGGAAAGTGTCTATGTTTGTGACTTTAGCGCTC 872
|||
Db 890 GAACAATGAAAGGAAACCTAAAAAGGAAAGTAACCTACAATATGAACCTAATGTG 949
|||
Qy 873 GGCAGCTATTTCCCTGGCAAG--TCTACACTCTCTGTCTGTGGGATTTGATCTTCGAGTCA 930
|||
Db 950 AACAGCTGTATACCTGGCCGCTATCTCCATCGGTGGAACCAAGAACTGAATCTTCGAGCTC 1009
|||
Qy 931 TCCTCTGTGTTCTCCAAATTCAT-AAACAGCTGTGGGATTTGCTATAGTTCCAGTTGCGGC 989
|||
Db 1010 TCCTCTCTATTTCAAATGCAATAAAGCGTTTGGAACTCCAGTGGTCTCTATCTCTGGC 1069
|||
Qy 990 AGTGATAGTTTCCATGATAGAAAGTGGACGCATCCAAATCTTCTCCATATTTTAGCTAT 1049
|||
Db 1070 TGTATCATTTTCCATGACAAAAATAAATGGGTAAAGGTGATTTCTATGATCTCTAGCTAT 1129
|||
Qy 1050 CTGCGGCTTCCTTCACTGCTCTATCAGCACTACTCGAAGAAAGAGTTGAATACTAG 1109
|||
Db 1130 TTGGGGTTTAACTTCCATGCTACCAAAAATATCTTGATGACAAAACTTTGAAGAAAAA 1189
|||
Qy 1110 CCACAAGTGTGTAGAGATCTTCATCTACCTGTTGAGGAAGTCCACAA 1162
|||
Db 1190 TCATGAATACAAACAGAAATCCCTGACCCACCAAGCAGAGAGTCA 1242
|||

CNS0A36P 1323 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GS1TFB64ZF05 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
BX826811 GI:42460388
HTC; GSLT cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1323)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1323)
CNS0A36P
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GS1TFB64ZF05 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
BX826811
GI:42460388
HTC; GSLT cDNA.
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
REFERENCE
AUTHORS
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1323)
Genoscope.
Direct Submission
TITLE
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefegenoscope.cns.fr
- Web : www.genoscope.cns.fr)
JOURNAL
The sequences are based on single pass reads.
Life technologies (a division of invitrogen) members carried out
full-length librairies construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
```

```
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
location/Qualifiers
1. .1323
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTFB64ZF05"
/tissue_type="Flowers and buds"
/plasmid="pCMVSPORT_6"
1. .1323
/gene="At4g18190"
ORIGIN
Query Match 26.2%; Score 312.8; DB 3; Length 1323;
Best Local Similarity 56.5%; Pred. No. 4.7e-78;
Matches 609; Conservative 0; Mismatches 457; Indels 12; Gaps 1;
Qy 53 AAGCAACCTTATAGATCATGAGGTGTAAGTGAATCATCATCATGAGTGTGCTCAAA 112
|||
Db 27 AATCAAGACTCCACTCAGTAAATGGCGACTGAAGGAAATTTTCAACAGAGAGAGAA 86
|||
Qy 113 CCAGAAACTATAAAGGTGCTTCTCTCCATATACGTAATCTTTGTCTCTTTTGCC 172
|||
Db 87 GTCAAGATCTCTTGAGGTAAAGTGTCTCTATGTCACCTCTCTCTAGCTGAG 146
|||
Qy 173 AGCCACTAGTACAATTTCTGGGTAGATTGTACTATGAAATGAGAGAAATAGACATATG 232
|||
Db 147 AGCAATAGCCACTCTCTTAGGTAGACTTTTACTACGAAAAGGCGTAAAGACATGCG 206
|||
Qy 233 TGGTAACACTTCTCAACTCATTTGGCTTCCCTGCTACTGTTCTGTTCCGCTCTTTTCTC 292
|||
Db 207 TCGAAACCTTGGTTGAGTGTAGGGTTTCTTTAACCTTCTCTGCTATTATTACTTAA 266
|||
Qy 293 GAATCAGGCAACCAAAATCAACAGATACAAATTTCACTAGTCCCTTCTTCAACACCC 352
|||
Db 267 AGCTGAGCGCTCAAGACTTAAACCACTTACCAAAAAAATCTCTCTCTCTTGACAC 326
|||
Qy 353 TTGCATCGGTTTACTTTGTGCATGGAAGTGTAGTGTCCGCTTATGTTATTTGTCGAG 412
|||
Db 327 TATCTTTAGTGATATTGGACTTTGGCTTGTCTTGTGGCTGGACATTTGATTTTGACTCAT 386
|||
Qy 413 TAGGGTTGCTCTACTTACCACTCTCTACTTTCTCCCTCATCTTGGCCTCACAGTTGGCCT 472
|||
Db 387 TTGGGCTACTTTACCTTCTCTCAACTTTCTCTTTGATCTCTGCGTCGCAATTTGGCTT 446
|||
Qy 473 TCACTGCTTTTCTCATATTTCTTAACTCGGAAAAGTTCACTCTTTGATAGTCAAGTT 532
|||
Db 447 TTAACGCGCTCTCTCTTACTTCTTAACCTACAAAAAATCAACACCAATTTATATCAATT 506
|||
Qy 533 CTTTGTCTTCTCCTACATGTAATCTCTGCTCTTCTTGTGGTCAACACTGATTCAGAA---- 588
|||
Db 507 CACTTGTCTCTTAAACCATATCTTACACTTCTTGTATTATCCAAACATGAACCAAAATCTC 566
|||
Qy 589 -----NACTCAACTAATGTATCTAGAGTATGATCGGGTTCATATGTACCA 640
|||
Db 567 CTTCTTCTTCTTCAAAAGTCCGAGCCAAAGTCCAAAGTATGATTAGATATCTTCCGCGG 626
|||
Qy 641 TCGGTGCTTCCGCTGGGATTTGGACTGTTACTATCTCTCGATACAAATGCTCTTCAGGAAAG 700
|||
Db 627 TCGGTAGCTCAGCTGTTATTCTCTGGTGTCTTTTAAACAGATTACGCGTACGAAAGA 686
|||
Qy 701 TTTTCAGAAAGCATACATCTCTAGCAGTACGAGACTTGGCACTTACAGTCTCTAGTTG 760
|||
Db 687 TTCAAAAGAAATACACATTTCAAGGCTATTTTAGACATATTTTCCGCTCTATGTTAG 746
|||
```

QY 761 CGAGTTGTGTAGTTCTCATAGGACTTTTTCAGAGTGGAGAGTGGGAAACTTTTGCCAAAGT 820
 |||||
 Db 747 CCACCTTGTGTAGTTGTGGTAGGACTTTTGGAGTGGTGGTGGGAAAAAGCTGAGTACAG 806
 |||||
 QY 821 AGATGAGAACTACAACTCGGAAAGTGCATATGTTTGTGACTTTTAGCCTCGGAGCTA 880
 |||||
 Db 807 AATGGAAGAGTTTCAACTAGGGAAGAGCTACATATTTTGAATAACATCGGTTCACGA 866
 |||||
 QY 881 TTTCCTGGCAAGTCTACACTCTCGTCTGTGGGATTTGATCTTCGAGTCACTCTCTGTGT 940
 |||||
 Db 867 TATCATGGCAAGCTTGTGATTTGGAAGTGTGGTTTGTGATTCGAGTTTTCATCGCTTT 926
 |||||
 QY 941 TCTCCAATTCCATAACAGCTGTGGGATTCCTATAGTTCAGTTCGGGAGGATAGTTT 1000
 |||||
 Db 927 TTTCCAATATCATAGAACTCTTTGTTTACAAGTTGTGCTGTCTTCTGTCTTATCTTCT 986
 |||||
 QY 1001 TCCATGATAGTGGAGGATCCAAATCTTCTCCATTATTTAGTATCTCGGGCTTCC 1060
 |||||
 Db 987 TCGTGATAGATAAGTGAATCAAGTTGGTTGCAATGTTTGGCCATCTCGGGATTG 1046
 |||||
 QY 1061 TTTCATTCTCTATCAGCACTACTCGACGAAAGAGTTGAATCTAGCCACACAAAG 1118
 |||||
 Db 1047 TTTCATTATGTTATCAGCATTTATGTCATGATAGAAAGCAAGAACCAAGAG 1104
 |||||

RESULT 13
 AU226302/C
 LOCUS AU226302 RAPL14 Arabidopsis thaliana CDNA clone RAPL14-23-A22 3',
 mRNA sequence.
 ACCESSION AU226302.1 GI:19740949
 VERSION AU226302.1
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 498)
 Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
 Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A.,
 Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
 Large scale analysis of Arabidopsis full-length cDNA
 Unpublished (2002)
 TITLE Contact: Motoaki Seki
 JOURNAL Plant Functional Genomics Research Group
 COMMENT RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rtc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
 and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
 al., submitted for publication) digested with BamHI and SalI. This
 clone is in a modified pBluescript vector. Please visit our web
 site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
 details.

FEATURES
 source
 1..498
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAPL14-23-A22"
 /tissue_type="root"
 /lab_host="DH10B"
 /clone_lib="RAPL14"
 /note="Site_1: BamHI; Site_2: SalI"

Query Match 25.1%; Score 299.6; DB 1; Length 498;

Best Local Similarity 98.7%; Pred. No. 2.1e-74;
 Matches 302; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 889 CAAGTCTACACTCCCTGGTCTTGTGGGATTTGATCTTCGAGTCACTCTCTGTGTTCTCCAAT 948
 |||||
 Db 498 CAGGCTCTACACTCTTGGTCTTGTGGGATTTGATCTTCGAGTCACTCTCTGTGTTCTCCAAT 439
 |||||
 QY 949 TCCATAACAGCTGTGGGATTCCTATAGTTCAGTTCGCGCAGTATAGTTCATGAT 1008
 |||||
 Db 438 TCCATAACAGCTGTGGGATTCCTATAGTTCAGTTCGCGCAGTATAGTTCATGAT 379
 |||||
 QY 1009 AGAATGGAGCGATCCAAATCTTCTCCATTATTTAGTATCTCGGGCTTCCTTCATTTC 1068
 |||||
 Db 378 AGAATGGAGCGATCCAAATCTTCTCCATTATTTAGTATCTCGGGCTTCCTTCATTTC 319
 |||||
 QY 1069 GTCTATCAGCACTACTCTCGAGAAAGAGTTGAATCTAGCCACACAAAGTCTGTAGGA 1128
 |||||
 Db 318 GTCTATCAGCACTACTCTCGAGAAAGAGTTGAATCTAGCCACACAAAGTCTGTAGGA 259
 |||||
 QY 1129 GATCTTCATCTACCTGTTGAGGAAGTCCACACAAATAGTGTGTGATCAAAAGCAT 1188
 |||||
 Db 258 GATCTTCATCTACCTGTTGAGGAAGTCCACACAAATAGTGTGTGATCAAAAGCAT 199
 |||||
 QY 1189 ATTTCC 1194
 |||||
 Db 198 ATTTCC 193

RESULT 14
 BZ967767
 LOCUS BZ967767
 DEFINITION PUDGUS6TD ZM_0.6.1.0 KB Zea mays genomic clone ZMMBTa184115,
 genomic survey sequence.
 ACCESSION BZ967767.1 GI:29185853
 VERSION BZ967767
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 922)
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Bennetzen,J.
 Maize Genomics Consortium
 Unpublished (2003)
 CONTACT: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TP
 Class: sheared ends.

FEATURES
 source
 1..922
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBTa184115"
 /clone_lib="ZM_0.6.1.0 KB"
 /note="Vector: pCE4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
 Cot selected genomic DNA library"

Query Match 20.1%; Score 239.6; DB 8; Length 922;
 Best Local Similarity 58.2%; Pred. No. 3.7e-57;
 Matches 421; Conservative 0; Mismatches 302; Indels 0; Gaps 0;

QY 376 CGACTGTAGTGTGGCTTATGCTTATTTGTCGAGTAGGGTCTCTACTTACAGTTC 435
 |||||
 |||||


```
Db 5 GGCCTCATCTTGGTCCGAGTGTGATGATATTCCTATGCTCTTTGTACCTCTCCCTGTA 64
Qy 436 TCTACTTTCTCCCTCATCTTGGCCCTCAGCTGGCTTCACTGCTTTTCTCATATTC 495
Db 65 TCACATATTCGCTCATTTGGCTAGTCAGCTGGCTTCAATGCTGCTTCTCATATGTC 124
Qy 496 CTTAACTCGCAAAAGTTCACCTCTTGTAGTGTGCTTCTTCTCTCTCTCCTCACCTGTATCC 555
Db 125 CTAATGCTCAAAAGTTCACCCCAATCAATTTCACTCAGTAATTCCTTACCTTTTCC 184
Qy 556 TCTGCTCTTCTGTGTGTCACACTGATTCAGAAACCTCAACTAATGTATCTAGAGTACAG 615
Db 185 GCTGGCTTCTTGGAGTTGACGAAGATTCCTCAGGCTACCAATGTTTATGCGTGGGAAG 244
Qy 616 TATGTGATCGGGTTCATATGTACCATCGGTGCTTCCGCTGGGATTTGAGCTTTACTATCT 675
Db 245 TACATATTTGGGTTTGGCATTTGACCTTAGGAGCTCGGCCACATATCACTAAATTTCTCTCT 304
Qy 676 CTGATACAAATGCTCTTCAGGAAAGTTTTCACGAAGCATATCTCAGCAGTCAACGGAC 735
Db 305 CTAATGCAAGTCGCANTTCGAAGGTTATTAAGAAGGAACTTTCTCAGTCTGTTGAAT 364
Qy 736 TTGGCCATTTACCAGTCTCTAGTTGCGAGTTGTGTAGTTCTCATAGGACTTTTTCGAAGT 795
Db 365 ATGCAGATATATACAGCACTAGTGGCAACAGTAGCTTCTTATCGGTTTATTTGCAAGC 424
Qy 796 GGAGGTGGGAACTTTCAGAGTGAGATGAGAACTACAACTCGGGAAGTGTCTATAT 855
Db 425 GCGAGTGGAAAGCTTTTAGAGGAGAGATGATGCTTTCAGCTCAGAGAGGTTGCTCTAT 484
Qy 856 GTTTTGACTTTAGCTTCGGCAGCTATTTCTTGGCAAGTCTACACTCCTGGTCTTGTGGGA 915
Db 485 GTGATGACACTTCTAAGGACTGCTGTATCTTGGCAGATAGCTTCGTTAGGAGTGGTGGT 544
Qy 916 TTGATCTTCGAGTCATCTCTGTGTCTTCCAAATTCATAACAGCTGTGGGATTCGCTATA 975
Db 545 TTAATCTTTGTGTGTCATCACTCTTTTCAAATGTGATAAGCACACTGGCTCTACCCATC 604
Qy 976 GTTCAGTTGGGCGAGTATAGTTTTCATGATAGATGGAGCATCCAAATCTTCTCC 1035
Db 605 ATTCGATTTTGTGTGATTTCTTCCACGACAGATGGATGGAGTGAAGTATTTGCT 664
Qy 1036 ATTAATTTAGCTATCTGGGCTTCTTCAATTCGTTCTATCAGCACTACCTCGACGAAAG 1095
Db 665 ATGTTGATGGCACTCGGGATTCGTTTCATATGGATATCAATTATGTCAGTGACAG 724
Qy 1096 AAG 1098
Db 725 AAG 727
```

```
RESULT 15
AU27210/c
LOCUS AU27210.1 RAFL14 Arabidopsis thaliana cdna clone RAFL14-88-B12 3',
DEFINITION AU27210 428 bp mRNA linear EST 23-APR-2002
RAFL14-88-B12 3',
mRNA sequence.
ACCESSION AU27210
VERSION AU27210.1 GI:19741857
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 428)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, P., Kawai, J.,
Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A.,
Muramatsu, M., Hayashizaki, F., and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
```

RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

FEATURES
source 1..428
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAFL14-88-B12"
 /tissue_type="root"
 /lab_host="DH10B"
 /clone_lib="RAFL14"
 /notes="Site_1: BamHI; Site_2: SalI"

ORIGIN
Query Match 19.9% Score 238; DB 1; Length 428;
Best Local Similarity 89.1%; Pred.No. 9.6e-57;
Matches 269; Conservative 0; Mismatches 30; Indels 3; Gaps 1;
Qy 896 ACACCTCCTCTTGTGGGATTCATCTTCGAGTCATCTCTGTGTCTCCAAATTCATAA 955
Db 428 ACACCATTTGGCGTCTGGGACTGATCTTTGAGTCATCTCTGTGTCTCCAAATTCATAA 369
Qy 956 CAGCTGTGGGATTCCTATAGTTCAGTTGGGCGAGTAGTTTCCATGATAGATGG 1015
Db 368 CTGCTGTGGGATTCCTATAGTTCAGTTGGGCGAGTAGTTTCCATGATAGATGG 309
Qy 1016 ACACATCCAAATCTTCTCCATTTATTTAGCTATCTCGGCTCTCTTCTGCTATC 1075
Db 308 ACGGTCAAAGATCTTCTCCATCAATTTAGCTATCTGGGATTCATTTCAATTTGCTATC 249
Qy 1076 AGCACTACCTCGACGAAAGAGTTGAATCTAGCCACACAAAGTCTGTAGGATCTTC 1135
Db 248 AGCACTACCTCGACGAAAGAGTTGAAGACTAGCCACACAAAGTCTGTAGGATCTTC 189
Qy 1136 AT---CTACCTGTGAGGAAGTTCACAAACATACAAAGTGTGTGATCAAGCATATTT 1192
Db 188 ATCTACTACCTGTGAGGAAGTTCACAAACATACATATGATGATCAAAACATATTT 129
Qy 1193 CC 1194
Db 128 CC 127

Search completed: November 1, 2004, 22:30:22
Job time : 4257.75 secs